

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Murphy, Brian R.  
Collins, Peter L.  
Durbin, Anna P.  
Skiadopoulos, Mario H.  
Tao, Tao
- (ii) TITLE OF INVENTION: PRODUCTION OF ATTENUATED PARAINFLUENZA  
VIRUS VACCINES FROM CLONED NUCLEOTIDE SEQUENCE
- (iii) NUMBER OF SEQUENCES: 74
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
(B) STREET: Two Embarcadero Center, 8th Floor  
(C) CITY: San Francisco  
(D) STATE: California  
(E) COUNTRY: USA  
(F) ZIP: 94111
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: 09/  
(B) FILING DATE: 22-MAY-1998  
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 60/059,385  
(B) FILING DATE: 19-SEP-1997
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 60/047,575  
(B) FILING DATE: 23-MAY-1997
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: King, Jeffrey J.  
(B) REGISTRATION NUMBER: 38,515  
(C) REFERENCE/DOCKET NUMBER: 17634-000320
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 206-467-9600  
(B) TELEFAX: 415-576-0300

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15669 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCTGGGTACC GGGCCCGTCG ACGCGTATAT AGTTCCTCCT TTCAGCAAAA AACCCCTCAA

GACCCGTTTA	GAGGCCCCAA	GGGGTTATGC	TACTGCAGGC	TCTCCCTTAG	CCATCCGAGT	120
GGACGTGCGT	CCTCCTTCGG	ATGCCCAGGT	CGGACCGCGA	GGAGGTGGAG	ATGCCATGCC	180
GACCCACCAA	ACAAGAGAAG	AAACTTGTCT	GGGAATATAA	ATTTAACTTT	AAATTAACTT	240
AGGATTAAAG	ACATTGACTA	GAAGGTCAAG	AAAAGGGAAC	TCTATAATTT	CAAAAATGTT	300
GAGCCTATTT	GATACATTTA	ATGCACGTAG	GCAAGAAAAC	ATAACAAAAT	CAGCCGGTGG	360
AGCTATCATT	CCTGGACAGA	AAAATACTGT	CTCTATATTC	GCCCTTGGAC	CGACAATAAC	420
TGATGATAAT	GAGAAAATGA	CATTAGCTCT	TCTATTTCTA	TCTCATTAC	TAGATAATGA	480
GAAACAACAT	GCACAAAGGG	CAGGGTTCTT	GGTGTCTTTA	TTGTCAATGG	CTTATGCCAA	540
TCCAGAGCTC	TACCTAACAA	CAAATGGAAG	TAATGCAGAT	GTCAAGTATG	TCATATACAT	600
GATTGAGAAA	GATCTAAAAC	GGCAAAAGTA	TGGAGGATTT	GTGGTTAAGA	CGAGAGAGAT	660
GATATATGAA	AAGACAACCTG	ATTGGATATT	TGGAAGTGAC	CTGGATTATG	ATCAGGAAAC	720
TATGTTGCAG	AACGGCAGGA	ACAATTCAAC	AATTGAAGAC	CTTGTCCACA	CATTTGGGTA	780
TCCATCATGT	TTAGGAGCTC	TTATAATACA	GATCTGGATA	GTTCTGGTCA	AAGCTATCAC	840
TAGTATCTCA	GGGTTAAGAA	AAGGCTTTTT	CACCCGATTG	GAAGCTTTCA	GACAAGATGG	900
AACAGTGCAG	GCAGGGCTGG	TATTGAGCGG	TGACACAGTG	GATCAGATTG	GGTCAATCAT	960
GCGGTCTCAA	CAGAGCTTGG	TAACTCTTAT	GGTTGAAACA	TTAATAACAA	TGAATACCAG	1020
CAGAAATGAC	CTCACAAACCA	TAGAAAAGAA	TATACAAATT	GTTGGCAACT	ACATAAGAGA	1080
TGCAGGTCTC	GCTTCATTCT	TCAATACAAT	CAGATATGGA	ATTGAGACCA	GAATGGCAGC	1140
TTTGACTCTA	TCCACTCTCA	GACCAGATAT	CAATAGATTA	AAAGCTTTGA	TGGAACTGTA	1200
TTTATCAAAG	GGACCACGCG	CTCCTTTCAT	CTGTATCCTC	AGAGATCCTA	TACATGGTGA	1260
GTTCGCACCA	GGCAACTATC	CTGCCATATG	GAGCTATGCA	ATGGGGGTGG	CAGTTGTACA	1320
AAATAGAGCC	ATGCAACAGT	ATGTGACGGG	AAGATCATAT	CTAGACATTG	ATATGTTCCA	1380
GCTAGGACAA	GCAGTAGCAC	GTGATGCCGA	AGCTCAAATG	AGCTCAACAC	TGGAAGATGA	1440
ACTTGAGAGT	ACACACGAAT	CTAAAGAAAG	CTTGAAGAGA	CATATAAGGA	ACATAAACAG	1500
TTCAGAGACA	TCTTTCCACA	AACCGACAGG	TGGATCAGCC	ATAGAGATGG	CAATAGATGA	1560
AGAGCCAGAA	CAATTCGAAC	ATAGAGCAGA	TCAAGAACAA	AATGGAGAAC	CTCAATCATC	1620
CATAATTCAA	TATGCCTGGG	CAGAAGGAAA	TAGAAGCGAT	GATCAGACTG	AGCAAGCTAC	1680
AGAATCTGAC	AATATCAAGA	CCGAACAACA	AAACATCAGA	GACAGACTAA	ACAAGAGACT	1740
CAACGACAAG	AAGAAACAAA	GCAGTCAACC	ACCCACTAAT	CCCACAAACA	GAACAAACCA	1800
GGACGAAATA	GATGATCTGT	TTAACGCATT	TGGAAGCAAC	TAATCGAATC	AACATTTTAA	1860
TCTAAATCAA	TAATAAATAA	GAAAACTTA	GGATTAAAGA	ATCCTATCAT	ACCGGAATAT	1920
AGGGTGGTAA	ATTTAGAGTC	TGCTTGAAAC	TCAATCAATA	GAGAGTTGAT	GGAAAGCGAT	1980
GCTAAAAACT	ATCAAATCAT	GGATTCTTGG	GAAGAGGAAT	CAAGAGATAA	ATCAACTAAT	2040
ATCTCCTCGG	CCCTCAACAT	CATTGAATTC	ATACTCAGCA	CCGACCCCCA	AGAAGACTTA	2100
TCGGAAAACG	ACACAATCAA	CACAAGAACC	CAGCAACTCA	GTGCCACCAT	CTGTCAACCA	2160

GAAATCAAAC	CAACAGAAAC	AAGTGAGAAA	GATAGTGGAT	CAACTGACAA	AAATAGACAG	2220
TCTGGGTCAT	CACACGAATG	TACAACAGAA	GCAAAAGATA	GAAATATTGA	TCAGGAAACT	2280
GTACAGAGAG	GACCTGGGAG	AAGAAGCAGC	TCAGATAGTA	GAGCTGAGAC	TGTGGTCTCT	2340
GGAGGAATCC	CCAGAAGCAT	CACAGATTCT	AAAAATGGAA	CCCAAACAC	GGAGGATATT	2400
GATCTCAATG	AAATTAGAAA	GATGGATAAG	GACTCTATTG	AGGGGAAAAT	GCGACAATCT	2460
GCAAATGTTC	CAAGCGAGAT	ATCAGGAAGT	GATGACATAT	TTACAACAGA	ACAAAGTAGA	2520
AACAGTGATC	ATGGAAGAAG	CCTGGAATCT	ATCAGTACAC	CTGATACAAG	ATCAATAAGT	2580
GTTGTTACTG	CTGCAACACC	AGATGATGAA	GAAGAAATAC	TAATGAAAAA	TAGTAGGACA	2640
AAGAAAAGTT	CTTCAACACA	TCAAGAAGAT	GACAAAAGAA	TTAAAAAAGG	GGGAAAAAGG	2700
AAAGACTGGT	TTAAGAAATC	AAAAGATACC	GACAACCAGA	TACCAACATC	AGACTACAGA	2760
TCCACATCAA	AAGGGCAGAA	GAAAATCTCA	AAGACAACAA	CCACCAACAC	CGACACAAAG	2820
GGGCAAACAG	AAATACAGAC	AGAATCATCA	GAAACACAAT	CCTCATCATG	GAATCTCATC	2880
ATCGACAACA	ACACCGACCG	GAACGAACAG	ACAAGCACAA	CTCCTCCAAC	AACAACCTCC	2940
AGATCAACTT	ATACAAAAGA	ATCGATCCGA	ACAACTCTG	AATCCAAACC	CAAGACACAA	3000
AAGACAAATG	GAAAGGAAAAG	GAAGGATACA	GAAGAGAGCA	ATCGATTTAC	AGAGAGGGCA	3060
ATTACTCTAT	TGCAGAATCT	TGGTGTAAAT	CAATCCACAT	CAAACTAGA	TTTATATCAA	3120
GACAAACGAG	TTGTATGTGT	AGCAAATGTA	CTAAACAATG	TAGATACTGC	ATCAAAGATA	3180
GATTTCTTGG	CAGGATTAGT	CATAGGGGTT	TCAATGGACA	ACGACACAAA	ATTAACACAG	3240
ATACAAAATG	AAATGCTAAA	CCTCAAAGCA	GATCTAAAGA	AAATGGACGA	ATCACATAGA	3300
AGATTGATAG	AAAATCAAAG	AGAACAACCTG	TCATTGATCA	CGTCACTAAT	TTCAAATCTC	3360
AAAATTATGA	CTGAGAGAGG	AGGAAAGAAA	GACCAAATG	AATCCAATGA	GAGAGTATCC	3420
ATGATCAAAA	CAAAATTGAA	AGAAGAAAAG	ATCAAGAAGA	CCAGGTTTGA	CCCACTTATG	3480
GAGGCACAAG	GCATTGACAA	GAATATACCC	GATCTATATC	GACATGCAGG	AGATACACTA	3540
GAGAACGATG	TACAAGTTAA	ATCAGAGATA	TTAAGTTCAT	ACAATGAGTC	AAATGCAACA	3600
AGACTAATAC	CCAAAAAAGT	GAGCAGTACA	ATGAGATCAC	TAGTTGCAGT	CATCAACAAC	3660
AGCAATCTCT	CACAAAGCAC	AAAACAATCA	TACATAAACG	AACTCAAACG	TTGCAAAAAT	3720
GATGAAGAAG	TATCTGAATT	AATGGACATG	TTCAATGAAG	ATGTCAACAA	TTGCCAATGA	3780
TCCAACAAAG	AAACGACACC	GAACAAACAG	ACAAGAAACA	ACAGTAGATC	AAAACCTGTC	3840
AACACACACA	AAATCAAGCA	GAATGAAACA	ACAGATATCA	ATCAATATAC	AAATAAGAAA	3900
AACTTAGGAT	TAAAGAATAA	ATTAATCCTT	GTCCAAAATG	AGTATAACTA	ACTCTGCAAT	3960
ATACACATTC	CCAGAATCAT	CATTCTCTGA	AAATGGTCAT	ATAGAACCAT	TACCACTCAA	4020
AGTCAATGAA	CAGAGGAAAG	CAGTACCCCA	CATTAGAGTT	GCCAAGATCG	GAAATCCACC	4080
AAAACACGGA	TCCCGGTATT	TAGATGTCTT	CTTACTCGGC	TTCTTCGAGA	TGGAACGAAT	4140
CAAAGACAAA	TACGGGAGTG	TGAATGATCT	CGACAGTGAC	CCGAGTTACA	AAGTTTGTGG	4200
CTCTGGATCA	TTACCAATCG	GATTGGCTAA	GTACACTGGG	AATGACCAGG	AATTGTTACA	4260

AGCCGCAACC	AAACTGGATA	TAGAAGTGAG	AAGAACAGTC	AAAGCGAAAG	AGATGGTTGT	4320
TTACACGGTA	CAAAATATAA	AACCAGAACT	GTACCCATGG	TCCAATAGAC	TAAGAAAAGG	4380
AATGCTGTTC	GATGCCAACA	AAGTTGCTCT	TGCTCCTCAA	TGTCTTCCAC	TAGATAGGAG	4440
CATAAAATTT	AGAGTAATCT	TCGTGAATTG	TACGGCAATT	GGATCAATAA	CCTTGTTCAA	4500
AATTCCTAAG	TCAATGGCAT	CACTATCTCT	ACCCAACACA	ATATCAATCA	ATCTGCAGGT	4560
ACACATAAAA	ACAGGGGTTC	AGACTGATTC	TAAAGGGATA	GTTCAAATTT	TGGATGAGAA	4620
AGGCGAAAAA	TCACTGAATT	TCATGGTCCA	TCTCGGATTG	ATCAAAAGAA	AAGTAGGCAG	4680
AATGTACTCT	GTTGAATACT	GTAAACAGAA	AATCGAGAAA	ATGAGATTGA	TATTTTCTTT	4740
AGGACTAGTT	GGAGGAATCA	GTCTTCATGT	CAATGCAACT	GGGTCCATAT	CAAAAACACT	4800
AGCAAGTCAG	CTGGTATTCA	AAAGAGAGAT	TTGTTATCCT	TTAATGGATC	TAAATCCGCA	4860
TCTCAATCTA	GTTATCTGGG	CTTCATCAGT	AGAGATTACA	AGAGTGGATG	CAATTTTCCA	4920
ACCTTCTTTA	CCTGGCGAGT	TCAGATACTA	TCCTAATATT	ATTGCAAAAG	GAGTTGGGAA	4980
AATCAAACAA	TGGAAC TAGT	AATCTCTATT	TTAGTCCGGA	CGTATCTATT	AAGCCGAAGC	5040
AAATAAGGA	TAATCAAAAA	CTTAGGACAA	AAGAGGTCAA	TACCAACAAC	TATTAGCAGT	5100
CACACTCGCA	AGAATAAGAG	AGAAGGGACC	AAAAAAGTCA	AATAGGAGAA	ATCAAAAACAA	5160
AAGGTACAGA	ACACCAGAAC	AACAAAATCA	AAACATCCAA	CTCACTCAAA	ACAAAAATTC	5220
CAAAAGAGAC	CGGCAACACA	ACAAGCACTG	AACACAATGC	CAACTTCAAT	ACTGCTAATT	5280
ATTACAACCA	TGATCATGGC	ATCTTCTGTC	CAAATAGATA	TCACAAAAC	ACAGCACGTA	5340
GGTGTATTGG	TCAACAGTCC	CAAAGGGATG	AAGATATCAC	AAAAC TTTGA	AACAAGATAT	5400
CTAATTTTGA	GCCTCATACC	AAAAATAGAA	GACTCTAACT	CTTGTGGTGA	CCAACAGATC	5460
AAGCAATACA	AGAAGTTATT	GGATAGACTG	ATCATCCCTT	TATATGATGG	ATTAAGATTA	5520
CAGAAAGATG	TGATAGTAAC	CAATCAAGAA	TCCAATGAAA	AACTGATCC	CAGAACAAAA	5580
CGATTCTTTG	GAGGGGTAAT	TGGAACCATT	GCTCTGGGAG	TAGCAACCTC	AGCACAAATT	5640
ACAGCGGCAG	TTGCTCTGGT	TGAAGCCAAG	CAGGCAAGAT	CAGACATCGA	AAAAC TCAAA	5700
GAAGCAATTA	GGGACACAAA	CAAAGCAGTG	CAGTCAGTTC	AGAGCTCCAT	AGGAAATTTA	5760
ATAGTAGCAA	TTAAATCAGT	CCAGGATTAT	GTTAACAAAG	AAATCGTGCC	ATCGATTGCG	5820
AGGCTAGGTT	GTGAAGCAGC	AGGACTTCAA	TTAGGAATTG	CATTAACACA	GCATTACTCA	5880
GAATTAACAA	ACATATTTGG	TGATAACATA	GGATCGTTAC	AAGAAAAAGG	AATAAAATTA	5940
CAAGGTATAG	CATCATTATA	CCGCACAAAT	ATCACAGAAA	TATTCACAAC	ATCAACAGTT	6000
GATAAATATG	ATATCTATGA	TCTGTTATTT	ACAGAATCAA	TAAAGGTGAG	AGTTATAGAT	6060
GTTGACTTGA	ATGATTACTC	AATCACCTC	CAAGTCAGAC	TCCCTTTATT	AACTAGGCTG	6120
CTGAACACTC	AGATCTACAA	AGTAGATTCC	ATATCATATA	ACATCCAAAA	CAGAGAATGG	6180
TATATCCCTC	TTCCAGCCA	TATCATGACG	AAAGGGGCAT	TTCTAGGTGG	AGCAGACGTC	6240
AAAGAATGTA	TAGAAGCATT	CAGCAGCTAT	ATATGCCCTT	CTGATCCAGG	ATTTGTATTA	6300
AACCATGAAA	TAGAGAGCTG	CTTATCAGGA	AACATATCCC	AATGTCCAAG	AACAACGGTC	6360

ACATCAGACA	TTGTTCCAAG	ATATGCATTT	GTCAATGGAG	GAGTGGTTGC	AAACTGTATA	6420
ACAACCACCT	GTACATGCAA	CGGAATTGGT	AATAGAATCA	ATCAACCACC	TGATCAAGGA	6480
GTAAAAATTA	TAACACATAA	AGAATGTAGT	ACAATAGGTA	TCAACGGAAT	GCTGTTCAAT	6540
ACAAATAAAG	AAGGAACTCT	TGCATTCTAT	ACACCAAATG	ATATAACACT	AAACAATTCT	6600
GTTGCACTTG	ATCCAATTGA	CATATCAATC	GAGCTCAACA	AGGCCAAATC	AGATCTAGAA	6660
GAATCAAAAG	AATGGATAAG	AAGGTCAAAT	CAAAACTAG	ATTCTATTGG	AAATTGGCAT	6720
CAATCTAGCA	CTACAATCAT	AATTATTTTG	ATAATGATCA	TTATATTGTT	TATAATTAAT	6780
ATAACGATAA	TTACAATTGC	AATTAAGTAT	TACAGAATTC	AAAAGAGAAA	TCGAGTGGAT	6840
CAAAATGACA	AGCCATATGT	ACTAACAAAC	AAATAACATA	TCTACAGATC	ATTAGATATT	6900
AAAATTATAA	AAAACCTTAGG	AGTAAAGTTA	CGCAATCCAA	CTCTACTCAT	ATAATTGAGG	6960
AAGGACCCAA	TAGACAAATC	CAAATTCGAG	ATGGAATACT	GGAAGCATAC	CAATCACGGA	7020
AAGGATGCTG	GTAATGAGCT	GGAGACGTCT	ATGGCTACTC	ATGGCAACAA	GCTCACTAAT	7080
AAGATAATAT	ACATATTATG	GACAATAATC	CTGGTGTTAT	TATCAATAGT	CTTCATCATA	7140
GTGCTAATTA	ATTCCATCAA	AAGTGAAAAG	GCCCACGAAT	CATTGCTGCA	AGACATAAAT	7200
AATGAGTTTA	TGGAAATTAC	AGAAAAGATC	CAAATGGCAT	CGGATAATAC	CAATGATCTA	7260
ATACAGTCAG	GAGTGAATAC	AAGGCTTCTT	ACAATTCAGA	GTCATGTCCA	GAATTACATA	7320
CCAATATCAT	TGACACAACA	GATGTCAGAT	CTTAGGAAAT	TCATTAGTGA	AATTACAATT	7380
AGAAATGATA	ATCAAGAAGT	GCTGCCACAA	AGAATAACAC	ATGATGTAGG	TATAAAAACCT	7440
TTAAATCCAG	ATGATTTTTG	GAGATGCACG	TCTGGTCTTC	CATCTTTAAT	GAAAACCTCCA	7500
AAAATAAGGT	TAATGCCAGG	GCCGGGATTA	TTAGCTATGC	CAACGACTGT	TGATGGCTGT	7560
GTTAGAACTC	CGTCTTTAGT	TATAAATGAT	CTGATTTATG	CTTATACCTC	AAATCTAATT	7620
ACTCGAGGTT	GTCAGGATAT	AGGAAAATCA	TATCAAGTCT	TACAGATAGG	GATAATAACT	7680
GTAAACTCAG	ACTTGGTACC	TGACTTAAAT	CCTAGGATCT	CTCATACCTT	TAACATAAAT	7740
GACAATAGGA	AGTCATGTTC	TCTAGCACTC	CTAAATATAG	ATGTATATCA	ACTGTGTTCA	7800
ACTCCCAAAG	TTGATGAAAG	ATCAGATTAT	GCATCATCAG	GCATAGAAGA	TATTGTACTT	7860
GATATTGTCA	ATTATGATGG	TTCAATCTCA	ACAACAAGAT	TTAAGAATAA	TAACATAAGC	7920
TTTGATCAAC	CATATGCTGC	ACTATACCCA	TCTGTTGGAC	CAGGGATATA	CTACAAAGGC	7980
AAAATAATAT	TTCTCGGGTA	TGGAGGTCTT	GAACATCCAA	TAAATGAGAA	TGTAATCTGC	8040
AACACAACCTG	GGTGCCCCGG	GAAAACACAG	AGAGACTGTA	ATCAAGCATC	TCATAGTACT	8100
TGGTTTTTCAG	ATAGGAGGAT	GGTCAACTCC	ATCATTGTTG	TTGACAAAGG	CTTAAACTCA	8160
ATTCCAAAAT	TGAAAGTATG	GACGATATCT	ATGCGACAAA	ATTACTGGGG	GTCAGAAGGA	8220
AGGTTACTTC	TACTAGGTAA	CAAGATCTAT	ATATATACAA	GATCTACAAG	TTGGCATAGC	8280
AAGTTACAAT	TAGGAATAAT	TGATATTACT	GATTACAGTG	ATATAAGGAT	AAAATGGACA	8340
TGGCATAATG	TGCTATCAAG	ACCAGGAAAC	AATGAATGTC	CATGGGGACA	TTCATGTCCA	8400
GATGGATGTA	TAACAGGAGT	ATATACTGAT	GCATATCCAC	TCAATCCCAC	AGGGAGCATT	8460

GTGTCATCTG	TCATATTAGA	CTCACAAAAA	TCGAGAGTGA	ACCCAGTCAT	AACTTACTCA	8520
ACAGCAACCG	AAAGAGTAAA	CGAGCTGGCC	ATCCTAAACA	GAACACTCTC	AGCTGGATAT	8580
ACAACAACAA	GCTGCATTAC	ACACTATAAC	AAAGGATATT	GTTTTTCATAT	AGTAGAAATA	8640
AATCATAAAA	GCTTAAACAC	ATTTCAACCC	ATGTTGTTCA	AAACAGAGAT	TCCAAAAAGC	8700
TGCAGTTAAT	CATAATTAAC	CATAATATGC	ATCAATCTAT	CTATAATACA	AGTATATGAT	8760
AAGTAATCAG	CAATCAGACA	ATAGACAAAA	GGGAAATATA	AAAACTTAG	GAGCAAAGCG	8820
TGCTCGGGAA	ATGGACACTG	AATCTAACAA	TGGCACTGTA	TCTGACATAC	TCTATCCTGA	8880
GTGTCACCTT	AACTCTCCTA	TCGTTAAAGG	TAAAATAGCA	CAATTACACA	CTATTATGAG	8940
TCTACCTCAG	CCTTATGATA	TGGATGACGA	CTCAATACTA	GTTATCACTA	GACAGAAAAT	9000
AAAACCTTAAT	AAATTGGATA	AAAGACAACG	ATCTATTAGA	AGATTAAAAT	TAATATTAAC	9060
TGAAAAAGTG	AATGACTTAG	GAAAATACAC	ATTTATCAGA	TATCCAGAAA	TGTCAAAAGA	9120
AATGTTCAAA	TTATATATAC	CTGGTATTAA	CAGTAAAGTG	ACTGAATTAT	TACTTAAAGC	9180
AGATAGAACA	TATAGTCAAA	TGACTGATGG	ATTAAGAGAT	CTATGGATTA	ATGTGCTATC	9240
AAAATTAGCC	TCAAAAAATG	ATGGAAGCAA	TTATGATCTT	AATGAAGAAA	TTAATAATAT	9300
ATCGAAAGTT	CACACAACCT	ATAAATCAGA	TAAATGGTAT	AATCCATTCA	AAACATGGTT	9360
TACTATCAAG	TATGATATGA	GAAGATTACA	AAAAGCTCGA	AATGAGATCA	CTTTTAATGT	9420
TGGGAAGGAT	TATAACTTGT	TAGAAGACCA	GAAGAATTTC	TTATTGATAC	ATCCAGAATT	9480
GGTTTTGATA	TTAGATAAAC	AAAACATATA	TGGTTATCTA	ATTACTCCTG	AATTAGTATT	9540
GATGTATTGT	GACGTAGTCG	AAGGCCGATG	GAATATAAGT	GCATGTGCTA	AGTTAGATCC	9600
AAAATTACAA	TCTATGTATC	AGAAAGGTAA	TAACCTGTGG	GAAGTGATAG	ATAAATTGTT	9660
TCCAATTATG	GGAGAAAAGA	CATTTGATGT	GATATCGTTA	TTAGAACCAC	TTGCATTATC	9720
CTTAATTCAA	ACTCATGATC	CTGTTAAACA	ACTAAGAGGA	GCTTTTTTTAA	ATCATGTGTT	9780
ATCCGAGATG	GAATTAATAT	TTGAATCTAG	AGAATCGATT	AAGGAATTTC	TGAGTGTAGA	9840
TTACATTGAT	AAAATTTTAG	ATATATTTAA	TAAGTCTACA	ATAGATGAAA	TAGCAGAGAT	9900
TTTCTCTTTT	TTTAGAACAT	TTGGGCATCC	TCCATTAGAA	GCTAGTATTG	CAGCAGAAAA	9960
GGTTAGAAAA	TATATGTATA	TTGGAAAAACA	ATTAAAATTT	GACACTATTA	ATAAATGTCA	10020
TGCTATCTTC	TGTACAATAA	TAATTAACGG	ATATAGAGAG	AGGCATGGTG	GACAGTGGCC	10080
TCCTGTGACA	TTACCTGATC	ATGCACACGA	ATTCATCATA	AATGCTTACG	GTTCAAACCTC	10140
TGCGATATCA	TATGAAAATG	CTGTTGATTA	TTACCAGAGC	TTTATAGGAA	TAAAATTCAA	10200
TAAATTCATA	GAGCCTCAGT	TAGATGAGGA	TTTGACAATT	TATATGAAAG	ATAAAGCATT	10260
ATCTCCAAAA	AAATCAAATT	GGGACACAGT	TTATCCTGCA	TCTAATTTAC	TGTACCGTAC	10320
TAACGCATCC	AACGAATCAC	GAAGATTAGT	TGAAGTATTT	ATAGCAGATA	GTAAATTTGA	10380
TCCTCATCAG	ATATTGGATT	ATGTAGAATC	TGGGGACTGG	TTAGATGATC	CAGAATTTAA	10440
TATTTCTTAT	AGTCTTAAAG	AAAAAGAGAT	CAAACAGGAA	GGTAGACTCT	TTGCAAAAAT	10500
GACATACAAA	ATGAGAGCTA	CACAAGTTTT	ATCAGAGACC	CTACTTGCAA	ATAACATAGG	10560

AAAATTCTTT	CAAGAAAATG	GGATGGTGAA	GGGAGAGATT	GAATTACTTA	AGAGATTAAC	10620
AACCATATCA	ATATCAGGAG	TTCCACGGTA	TAATGAAGTG	TACAATAATT	CTAAAAGCCA	10680
TACAGATGAC	CTTAAACCT	ACAATAAAAT	AAGTAATCTT	AATTTGTCTT	CTAATCAGAA	10740
ATCAAAGAAA	TTTGAATTCA	AGTCAACGGA	TATCTACAAT	GATGGATACG	AGACTGTGAG	10800
CTGTTTCCTA	ACAACAGATC	TCAAAAAATA	CTGTCTTAAT	TGGAGATATG	AATCAACAGC	10860
TCTATTTGGA	GAAACTTGCA	ACCAAATATT	TGGATTAAAT	AAATTGTTTA	ATTGGTTACA	10920
CCCTCGTCTT	GAAGGAAGTA	CAATCTATGT	AGGTGATCCT	TACTGTCCTC	CATCAGATAA	10980
AGAACATATA	TCATTAGAGG	ATCACCCCTGA	TTCTGGTTTT	TACGTTCATA	ACCCAAGAGG	11040
GGGTATAGAA	GGATTTTGTC	AAAAATTATG	GACACTCATA	TCTATAAGTG	CAATACATCT	11100
AGCAGCTGTT	AGAATAGGCG	TGAGGGTGAC	TGCAATGGTT	CAAGGAGACA	ATCAAGCTAT	11160
AGCTGTAACC	ACAAGAGTAC	CCAACAATTA	TGACTACAGA	GTTAAGAAGG	AGATAGTTTA	11220
TAAAGATGTA	GTGAGATTTT	TTGATTCATT	AAGAGAAGTG	ATGGATGATC	TAGGTCATGA	11280
ACTTAAATTA	AATGAAACGA	TTATAAGTAG	CAAGATGTTC	ATATATAGCA	AAAGAATCTA	11340
TTATGATGGG	AGAATTCTTC	CTCAAGCTCT	AAAAGCATT	TCTAGATGTG	TCTTCTGGTC	11400
AGAGACAGTA	ATAGACGAAA	CAAGATCAGC	ATCTTCAAAT	TTGGCAACAT	CATTTGCAAA	11460
AGCAATTGAG	AATGGTTATT	CACCTGTTCT	AGGATATGCA	TGCTCAATTT	TTAAGAATAT	11520
TCAACAATA	TATATTGCCC	TTGGGATGAA	TATCAATCCA	ACTATAACAC	AGAATATCAG	11580
AGATCAGTAT	TTTAGGAATC	CAAATTGGAT	GCAATATGCC	TCTTTAATAC	CTGCTAGTGT	11640
TGGGGGATTC	AATTACATGG	CCATGTCAAG	ATGTTTTGTA	AGGAATATTG	GTGATCCATC	11700
AGTTGCCGCA	TTGGCTGATA	TTAAAAGATT	TATTAAGGCG	AATCTATTAG	ACCGAAGTGT	11760
TCTTTATAGG	ATTATGAATC	AAGAACCAGG	TGAGTCATCT	TTTTTGGACT	GGGCTTCAGA	11820
TCCATATTCA	TGCAATTTAC	CACAATCTCA	AAATATAACC	ACCATGATAA	AAAATATAAC	11880
AGCAAGGAAT	GTATTACAAG	ATTCACCAA	TCCATTATTA	TCTGGATTAT	TCACAAATAC	11940
AATGATAGAA	GAAGATGAAG	AATTAGCTGA	GTTCCTGATG	GACAGGAAGG	TAATTCTCCC	12000
TAGAGTTGCA	CATGATATTC	TAGATAATTC	TCTCACAGGA	ATTAGAAATG	CCATAGCTGG	12060
AATGTTAGAT	ACGACAAAAT	CACTAATTCG	GGTTGGCATA	AATAGAGGAG	GACTGACATA	12120
TAGTTTGTTG	AGGAAAATCA	GTAATTACGA	TCTAGTACAA	TATGAAACAC	TAAGTAGGAC	12180
TTTGCGACTA	ATTGTAAGTG	ATAAAATCAA	GTATGAAGAT	ATGTGTTCCG	TAGACCTTGC	12240
CATAGCATTG	CGACAAAAGA	TGTGGATTCA	TTTATCAGGA	GGAAGGATGA	TAAGTGACT	12300
TGAAACGCCT	GACCCATTAG	AATTACTATC	TGGGGTAGTA	ATAACAGGAT	CAGAACATTG	12360
TAAATATGT	TATTCTTCAG	ATGGCACAAA	CCCATATACT	TGGATGTATT	TACCCGTAA	12420
TATCAAAATA	GGATCAGCAG	AAACAGGTAT	ATCGTCATTA	AGAGTTCCTT	ATTTTGGATC	12480
AGTCACTGAT	GAAAGATCTG	AAGCACAATT	AGGATATATC	AAGAATCTTA	GTAAACCTGC	12540
AAAAGCCGCA	ATAAGAATAG	CAATGATATA	TACATGGGCA	TTTGGTAATG	ATGAGATATC	12600
TTGGATGGAA	GCCTCACAGA	TAGCACAAAC	ACGTGCAAAT	TTTACACTAG	ATAGTCTCAA	12660

AATTTTAACA	CCGGTAGCTA	CATCAACAAA	TTTATCACAC	AGATTAAAGG	ATACTGCAAC	12720
TCAGATGAAA	TTCTCCAGTA	CATCATTGAT	CAGAGTCAGC	AGATTCATAA	CAATGTCCAA	12780
TGATAACATG	TCTATCAAAG	AAGCTAATGA	AACCAAAGAT	ACTAATCTTA	TTTATCAACA	12840
AATAATGTTA	ACAGGATTAA	GTGTTTTCGA	ATATTTATTT	AGATTAAAAG	AAACCACAGG	12900
ACACAACCTT	ATAGTTATGC	ATCTGCACAT	AGAAGATGAG	TGTTGTATTA	AAGAAAGTTT	12960
TAATGATGAA	CATATTAATC	CAGAGTCTAC	ATTAGAATTA	ATTCGATATC	CTGAAAGTAA	13020
TGAATTTATT	TATGATAAAG	ACCCACTCAA	AGATGTGGAC	TTATCAAAAC	TTATGGTTAT	13080
TAAAGACCAT	TCTTACACAA	TTGATATGAA	TTATTGGGAT	GATACTGACA	TCATACATGC	13140
AATTTCAATA	TGTA CTGCAA	TTACAATAGC	AGATACTATG	TCACAATTAG	ATCGAGATAA	13200
TTTAAAAGAG	ATAATAGTTA	TTGCAAATGA	TGATGATATT	AATAGCTTAA	TCACTGAATT	13260
TTTGACTCTT	GACATACTTG	TATTTCTCAA	GACATTTGGT	GGATTATTAG	TAAATCAATT	13320
TGCATACACT	CTTTATAGTC	TAAAAATAGA	AGGTAGGGAT	CTCATTTGGG	ATTATATAAT	13380
GAGAACACTG	AGAGATACTT	CCCATTCAAT	ATTAAAAGTA	TTATCTAATG	CATTATCTCA	13440
TCCTAAAGTA	TTCAAGAGGT	TCTGGGATTG	TGGAGTTTTA	AACCCTATTT	ATGGTCCTAA	13500
TACTGCTAGT	CAAGACCAGA	TAAAACTTGC	CCTATCTATA	TGTGAATATT	CACTAGATCT	13560
ATTTATGAGA	GAATGGTTGA	ATGGTGTATC	ACTTGAAATA	TACATTTGTG	ACAGCGATAT	13620
GGAAGTTGCA	AATGATAGGA	AACAAGCCTT	TATTTCTAGA	CACCTTTCAT	TTGTTTGTTG	13680
TTTAGCAGAA	ATTGCATCTT	TCGGACCTAA	CCTGTTAAAC	TTAACATACT	TGGAGAGACT	13740
TGATCTATTG	AAACAATATC	TTGAATTAAA	TATTAAAGAA	GACCCTACTC	TTAAATATGT	13800
ACAAATATCT	GGATTATTAA	TTAAATCGTT	CCCATCAACT	GTAACATACG	TAAGAAAGAC	13860
TGCAATCAAA	TATCTAAGGA	TTCGCGGTAT	TAGTCCACCT	GAGGTAATTG	ATGATTGGGA	13920
TCCGGTAGAA	GATGAAAATA	TGCTGGATAA	CATTGTCAAA	ACTATAAATG	ATAACTGTAA	13980
TAAAGATAAT	AAAGGGAATA	AAATTAACAA	TTTCTGGGGA	CTAGCACTTA	AGAACTATCA	14040
AGTCCTTAAA	ATCAGATCTA	TAACAAGTGA	TTCTGATGAT	AATGATAGAC	TAGATGCTAA	14100
TACAAGTGGT	TTGACACTTC	CTCAAGGAGG	GAATTATCTA	TCGCATCAAT	TGAGATTATT	14160
CGGAATCAAC	AGCACTAGTT	GTCTGAAAGC	TCTTGAGTTA	TCACAAATTT	TAATGAAGGA	14220
AGTCAATAAA	GACAAGGACA	GGCTCTTCCT	GGGAGAAGGA	GCAGGAGCTA	TGCTAGCATG	14280
TTATGATGCC	ACATTAGGAC	CTGCAGTTAA	TTATTATAAT	TCAGGTTTGA	ATATAACAGA	14340
TGTAATTGGT	CAACGAGAAT	TGAAAATATT	TCCTTCAGAG	GTATCATTAG	TAGGTAAAAA	14400
ATTAGGAAAT	GTGACACAGA	TTCTTAACAG	GGTAAAAGTA	CTGTTCAATG	GGAATCCTAA	14460
TTCAACATGG	ATAGGAAATA	TGGAATGTGA	GAGCTTAATA	TGGAGTGAAT	TAAATGATAA	14520
GTCCATTGGA	TTAGTACATT	GTGATATGGA	AGGAGCTATC	GGTAAATCAG	AAGAACTGT	14580
TCTACATGAA	CATTATAGTG	TTATAAGAAT	TACATACTTG	ATTGGGGATG	ATGATGTTGT	14640
TTTAGTTTCC	AAAATTATAC	CTACAATCAC	TCCGAATTGG	TCTAGAATAC	TTTATCTATA	14700
TAAATTATAT	TGGAAAGATG	TAAGTATAAT	ATCACTCAAA	ACTTCTAATC	CTGCATCAAC	14760



AGAATTATAT CTAATTTCTGA AAGATGCATA TTGTACTATA ATGGAACCTA GTGAAATTGT 14820  
 TTTATCAAAA CTTAAAAGAT TGTCACTCTT GGAAGAAAAT AATCTATTAA AATGGATCAT 14880  
 TTTATCAAAG AAGAGGAATA ATGAATGGTT ACATCATGAA ATCAAAGAAG GAGAAAGAGA 14940  
 TTATGGAATC ATGAGACCAT ATCATATGGC ACTACAAATC TTTGGATTTC AAATCAATTT 15000  
 AAATCATCTG GCGAAAGAAT TTTTATCAAC CCCAGATCTG ACTAATATCA ACAATATAAT 15060  
 CCAAAGTTTT CAGCGAACAA TAAAGGATGT TTTATTTGAA TGGATTAATA TAACTCATGA 15120  
 TGATAAGAGA CATAAATTAG GCGGAAGATA TAACATATTC CCACTGAAAA ATAAGGGAAA 15180  
 GTTAAGACTG CTATCGAGAA GACTAGTATT AAGTTGGATT TCATTATCAT TATCGACTCG 15240  
 ATTACTTACA GGTGCTTTTC CTGATGAAAA ATTTGAACAT AGAGCACAGA CTGGATATGT 15300  
 ATCATTAGCT GATACTGATT TAGAATCATT AAAGTTATTG TCGAAAAACA TCATTAAGAA 15360  
 TTACAGAGAG TGTATAGGAT CAATATCATA TTGGTTTCTA ACCAAAGAAG TTAAAATACT 15420  
 TATGAAATTG ATCGGTGGTG CTAAATTATT AGGAATTCCC AGACAATATA AAGAACCCGA 15480  
 AGACCAGTTA TTAGAAAAC TACAATCAACA TGATGAATTT GATATCGATT AAAACATAAA 15540  
 TACAATGAAG ATATATCCTA ACCTTTATCT TTAAGCCTAG GAATAGACAA AAAGTAAGAA 15600  
 AAACATGTAA TATATATATA CCAAACAGAG TTCTTCTCTT GTTTGGTTAT AGTGAGTCGT 15660  
 ATTACAATC 15669

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AATACGACTC ACTATAACCA AACAAGAGAA C 31

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCAAGTACTA TGAGATGCTT CATT 24

## (2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 base pairs  
 (B) TYPE: nucleic acid

(ii) MOLECULE TYPE: cDNA

CCCTATAATT TCAACATGTT GAGCCTATTT G

31

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

GATTAAATG TTGGTCGACT TAGTTGCTTC C

31

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 38 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

CCATAGAGAG TCCATGGAAA GCGACGCTAA AAACCTATC

38

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 37 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

CGGTGTCGTT TCTTTGTCGA CTCATTGGCA ATTGTTG

37

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 42 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

42

(ii) MOLECULE TYPE: cDNA

28

(ii) MOLECULE TYPE: cDNA

30

(ii) MOLECULE TYPE: cDNA

29

(ii) MOLECULE TYPE: cDNA

43

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 35 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

35

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 15660 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TAATACGACT	CACTATAACC	AAACAAGAGA	AGAAACTTGT	CTGGGAATAT	AAATTTAACT	60
TTAAATTAACT	TTAGGATTAA	AGACATTGAC	TAGAAGGTCA	AGAAAAGGGA	ACTCTATAAT	120
TTCAAAAATG	TTGAGCCTAT	TTGATACATT	TAATGCACGT	AGGCAAGAAA	ACATAACAAA	180
ATCAGCCGGT	GGAGCTATCA	TTCTTGACAC	GAAAAATACT	GTCTCTATAT	TCGCCCTTGG	240
ACCGACAATA	ACTGATGATA	ATGAGAAAAT	GACATTAGCT	CTTCTATTTC	TATCTCATTC	300
ACTAGATAAT	GAGAAACAAC	ATGCACAAAG	GGCAGGGTTC	TTGGTGTCTT	TATTGTCAAT	360
GGCTTATGCC	AATCCAGAGC	TCTACCTAAC	AACAAATGGA	AGTAATGCAG	ATGTCAAGTA	420
TGTCATATAC	ATGATTGAGA	AAGATCTAAA	ACGGCAAAAG	TATGGAGGAT	TTGTGGTTAA	480
GACGAGAGAG	ATGATATATG	AAAAGACAAC	TGATTGGATA	TTTGAAGTGA	ACCTGGATTA	540
TGATCAGGAA	ACTATGTTGC	AGAACGGCAG	GAACAATTCA	ACAATTGAAG	ACCTTGTTCA	600
CACATTTGGG	TATCCATCAT	GTTTAGGAGC	TCTTATAATA	CAGATCTGGA	TAGTTCCTGGT	660
CAAAGCTATC	ACTAGTATCT	CAGGGTTAAG	AAAAGGCTTT	TTCACCCGAT	TGGAAGCTTT	720
CAGACAAGAT	GGAACAGTGC	AGGCAGGGCT	GGTATTGAGC	GGTGACACAG	TGGATCAGAT	780
TGGGTCAATC	ATGCGGTCTC	AACAGAGCTT	GGTAACTCTT	ATGGTTGAAA	CATTAATAAC	840
AATGAATACC	AGCAGAAATG	ACCTCACAAC	CATAGAAAAG	AATATACAAA	TTGTTGGCAA	900
CTACATAAGA	GATGCAGGTC	TCGCTTCATT	CTTCAATACA	ATCAGATATG	GAATTGAGAC	960
CAGAATGGCA	GCTTTGACTC	TATCCACTCT	CAGACCAGAT	ATCAATAGAT	TAAAAGCTTT	1020
GATGGAAGTG	TATTTATCAA	AGGGACCACG	CGCTCCTTTC	ATCTGTATCC	TCAGAGATCC	1080
TATACATGGT	GAGTTCGCAC	CAGGCAACTA	TCCTGCCATA	TGGAGCTATG	CAATGGGGGT	1140
GGCAGTTGTA	CAAAATAGAG	CCATGCAACA	GTATGTGACG	GGAAGATCAT	ATCTAGACAT	1200

TGATATGTTT	CAGCTAGGAC	AAGCAGTAGC	ACGTGATGCC	GAAGCTCAAA	TGAGCTCAAC	1260
ACTGGAAGAT	GAAGTTGGAG	TGACACACGA	ATCTAAAGAA	AGCTTGAAGA	GACATATAAG	1320
GAACATAAAC	AGTTCAGAGA	CATCTTTCCA	CAAACCGACA	GGTGGATCAG	CCATAGAGAT	1380
GGCAATAGAT	GAAGAGCCAG	AACAATTCGA	ACATAGAGCA	GATCAAGAAC	AAAATGGAGA	1440
ACCTCAATCA	TCCATAATTC	AATATGCCTG	GGCAGAAGGA	AATAGAAGCG	ATGATCAGAC	1500
TGAGCAAGCT	ACAGAATCTG	ACAATATCAA	GACCGAACAA	CAAAACATCA	GAGACAGACT	1560
AAACAAGAGA	CTCAACGACA	AGAAGAAACA	AAGCAGTCAA	CCACCCACTA	ATCCCACAAA	1620
CAGAACAAAC	CAGGACGAAA	TAGATGATCT	GTTTAACGCA	TTTGGAAGCA	ACTAATCGAA	1680
TCAACATTTT	AATCTAAATC	AATAATAAAT	AAGAAAAACT	TAGGATTAAA	GAATCCTATC	1740
ATACCGGAAT	ATAGGGTGGT	AAATTTAGAG	TCTGCTTGAA	ACTCAATCAA	TAGAGAGTTG	1800
ATGGAAAGCG	ATGCTAAAAA	CTATCAAATC	ATGGATTCTT	GGGAAGAGGA	ATCAAGAGAT	1860
AAATCAACTA	ATATCTCCTC	GGCCCTCAAC	ATCATTGAAT	TCATACTCAG	CACCGACCCC	1920
CAAGAAGACT	TATCGGAAAA	CGACACAATC	AACACAAGAA	CCCAGCAACT	CAGTGCCACC	1980
ATCTGTCAAC	CAGAAATCAA	ACCAACAGAA	ACAAGTGAGA	AAGATAGTGG	ATCAACTGAC	2040
AAAAATAGAC	AGTCTGGGTC	ATCACACGAA	TGTACAACAG	AAGCAAAAGA	TAGAAATATT	2100
GATCAGGAAA	CTGTACAGAG	AGGACCTGGG	AGAAGAAGCA	GCTCAGATAG	TAGAGCTGAG	2160
ACTGTGGTCT	CTGGAGGAAT	CCCCAGAAGC	ATCACAGATT	CTAAAAATGG	AACCCAAAAC	2220
ACGGAGGATA	TTGATCTCAA	TGAAATTAGA	AAGATGGATA	AGGACTCTAT	TGAGGGGAAA	2280
ATGCGACAAT	CTGCAAATGT	TCCAAGCGAG	ATATCAGGAA	GTGATGACAT	ATTTACAACA	2340
GAACAAAGTA	GAAACAGTGA	TCATGGAAGA	AGCCTGGAAT	CTATCAGTAC	ACCTGATACA	2400
AGATCAATAA	GTGTTGTTAC	TGCTGCAACA	CCAGATGATG	AAGAAGAAAT	ACTAATGAAA	2460
AATAGTAGGA	CAAAGAAAAA	TTCTTCAACA	CATCAAGAAG	ATGACAAAAG	AATTAAAAAA	2520
GGGGGAAAAG	GGAAAGACTG	GTTTAAGAAA	TCAAAAGATA	CCGACAACCA	GATACCAACA	2580
TCAGACTACA	GATCCACATC	AAAAGGGCAG	AAGAAAATCT	CAAAGACAAC	AACCACCAAC	2640
ACCGACACAA	AGGGGCAAAC	AGAAATACAG	ACAGAATCAT	CAGAAACACA	ATCCTCATCA	2700
TGGAATCTCA	TCATCGACAA	CAACACCGAC	CGGAACGAAC	AGACAAGCAC	AACTCCTCCA	2760
ACAACAACCT	CCAGATCAAC	TTATACAAAA	GAATCGATCC	GAACAAACTC	TGAATCCAAA	2820
CCCAAGACAC	AAAAGACAAA	TGGAAAGGAA	AGGAAGGATA	CAGAAGAGAG	CAATCGATTT	2880
ACAGAGAGGG	CAATTACTCT	ATTGCAGAAT	CTTGGTGTA	TTCAATCCAC	ATCAAAACTA	2940
GATTTATATC	AAGACAAAAC	AGTTGTATGT	GTAGCAAATG	TACTAAACAA	TGTAGATACT	3000
GCATCAAAGA	TAGATTTCCCT	GGCAGGATTA	GTCATAGGGG	TTTCAATGGA	CAACGACACA	3060
AAATTAACAC	AGATACAAAA	TGAAATGCTA	AACCTCAAAG	CAGATCTAAA	GAAAATGGAC	3120
GAATCACATA	GAAGATTGAT	AGAAAATCAA	AGAGAACAAC	TGTCATTGAT	CACGTCACCTA	3180
ATTTCAAATC	TCAAAATTAT	GACTGAGAGA	GGAGGAAAGA	AAGACCAAAA	TGAATCCAAT	3240
GAGAGAGTAT	CCATGATCAA	AACAAAATTG	AAAGAAGAAA	AGATCAAGAA	GACCAGGTTT	3300

GACCCACTTA	TGGAGGCACA	AGGCATTGAC	AAGAAATATAC	CCGATCTATA	TCGACATGCA	3360
GGAGATACAC	TAGAGAACGA	TGTACAAGTT	AAATCAGAGA	TATTAAGTTC	ATACAATGAG	3420
TCAAATGCAA	CAAGACTAAT	ACCCAAAAAA	GTGAGCAGTA	CAATGAGATC	ACTAGTTGCA	3480
GTCATCAACA	ACAGCAATCT	CTCACAAAGC	ACAAAACAAT	CATACATAAA	CGAACTCAAA	3540
CGTTGCAAAA	ATGATGAAGA	AGTATCTGAA	TTAATGGACA	TGTTCAATGA	AGATGTCAAC	3600
AATTGCCAAT	GATCCAACAA	AGAAACGACA	CCGAACAAAC	AGACAAGAAA	CAACAGTAGA	3660
TCAAAACCTG	TCAACACACA	CAAAATCAAG	CAGAATGAAA	CAACAGATAT	CAATCAATAT	3720
ACAAATAAGA	AAAACCTTAGG	ATTAAAGAAT	AAATTAATCC	TTGTCCAAAA	TGAGTATAAC	3780
TAACTCTGCA	ATATACACAT	TCCCAGAATC	ATCATTCTCT	GAAAATGGTC	ATATAGAACC	3840
ATTACCACTC	AAAGTCAATG	AACAGAGGAA	AGCAGTACCC	CACATTAGAG	TTGCCAAGAT	3900
CGGAAATCCA	CCAAAACACG	GATCCC GGTA	TTTAGATGTC	TTCTTACTCG	GCTTCTTCGA	3960
GATGGAACGA	ATCAAAGACA	AATACGGGAG	TGTGAATGAT	CTCGACAGTG	ACCCGAGTTA	4020
CAAAGTTTGT	GGCTCTGGAT	CATTACCAAT	CGGATTGGCT	AAGTACACTG	GGAATGACCA	4080
GGAATTGTTA	CAAGCCGCAA	CCAAACTGGA	TATAGAAGTG	AGAAGAACAG	TCAAAGCGAA	4140
AGAGATGGTT	GTTTACACGG	TACAAAATAT	AAAACCAGAA	CTGTACCCAT	GGTCCAATAG	4200
ACTAAGAAAA	GGAATGCTGT	TCGATGCCAA	CAAAGTTGCT	CTTGCTCCTC	AATGTCTTCC	4260
ACTAGATAGG	AGCATAAAAT	TTAGAGTAAT	CTTCGTGAAT	TGTACGGCAA	TTGGATCAAT	4320
AACCTTGTTT	AAAATTCCTA	AGTCAATGGC	ATCACTATCT	CTACCCAACA	CAATATCAAT	4380
CAATCTGCAG	GTACACATAA	AAACAGGGGT	TCAGACTGAT	TCTAAAGGGA	TAGTTCAAAT	4440
TTTGGATGAG	AAAGGCGAAA	AATCACTGAA	TTTCATGGTC	CATCTCGGAT	TGATCAAAAG	4500
AAAAGTAGGC	AGAATGTACT	CTGTTGAATA	CTGTAAACAG	AAAATCGAGA	AAATGAGATT	4560
GATATTTTCT	TTAGGACTAG	TTGGAGGAAT	CAGTCTTCAT	GTCAATGCAA	CTGGGTCCAT	4620
ATCAAAAACA	CTAGCAAGTC	AGCTGGTATT	CAAAGAGAG	ATTTGTTATC	CTTTAATGGA	4680
TCTAAATCCG	CATCTCAATC	TAGTTATCTG	GGCTTCATCA	GTAGAGATTA	CAAGAGTGGA	4740
TGCAATTTTCT	CAACCTTCTT	TACCTGGCGA	G TTCAGATAC	TATCCTAATA	TTATTGCAAA	4800
AGGAGTTGGG	AAAATCAAAC	AATGGA ACTA	GTAATCTCTA	TTTTAGTCCG	GACGTATCTA	4860
TTAAGCCGAA	GCAAATAAAG	GATAATCAAA	AACTTAGGAC	AAAAGAGGTC	AATACCAACA	4920
ACTATTAGCA	GTCACACTCG	CAAGAATAAG	AGAGAAGGGA	CCAAAAAAGT	CAAATAGGAG	4980
AAATCAAAAC	AAAAGGTACA	GAACACCAGA	ACAACAAAT	CAAAACATCC	AACTCACTCA	5040
AAACAAAAAT	TCCAAAAGAG	ACCGGCAACA	CAACAAGCAC	TGAACACAAT	GCCAACTTCA	5100
ATACTGCTAA	TTATTACAAC	CATGATCATG	GCATCTTTCT	GCCAAATAGA	TATCACAAAA	5160
CTACAGCACG	TAGGTGTATT	GGTCAACAGT	CCCAAAGGGA	TGAAGATATC	ACAAA ACTTT	5220
GAAACAAGAT	ATCTAATTTT	GAGCCTCATA	CCAAAAATAG	AAGACTCTAA	CTCTTGTTGGT	5280
GACCAACAGA	TCAAGCAATA	CAAGAAGTTA	TTGGATAGAC	TGATCATCCC	TTTATATGAT	5340
GGATTAAGAT	TACAGAAAGA	TGTGATAGTA	ACCAATCAAG	AATCCAATGA	AAACACTGAT	5400

CCCAGAACAA	AACGATTCTT	TGGAGGGGTA	ATTGGAACCA	TTGCTCTGGG	AGTAGCAACC	5460
TCAGCACAAA	TTACAGCGGC	AGTTGCTCTG	GTTGAAGCCA	AGCAGGCAAG	ATCAGACATC	5520
GAAAAACTCA	AAGAAGCAAT	TAGGGACACA	AACAAAGCAG	TGCAGTCAGT	TCAGAGCTCC	5580
ATAGGAAATT	TAATAGTAGC	AATTAAATCA	GTCCAGGATT	ATGTTAACAA	AGAAATCGTG	5640
CCATCGATTG	CGAGGCTAGG	TTGTGAAGCA	GCAGGACTTC	AATTAGGAAT	TGCATTAACA	5700
CAGCATTACT	CAGAATTAAC	AAACATATTT	GGTGATAACA	TAGGATCGTT	ACAAGAAAAA	5760
GGAATAAAAT	TACAAGGTAT	AGCATCATTA	TACCGCACAA	ATATCACAGA	AATATTCACA	5820
ACATCAACAG	TTGATAAATA	TGATATCTAT	GATCTGTTAT	TTACAGAATC	AATAAAGGTG	5880
AGAGTTATAG	ATGTTGACTT	GAATGATTAC	TCAATCACCC	TCCAAGTCAG	ACTCCCTTTA	5940
TTAACTAGGC	TGCTGAACAC	TCAGATCTAC	AAAGTAGATT	CCATATCATA	TAACATCCAA	6000
AACAGAGAAT	GGTATATCCC	TCTTCCCAGC	CATATCATGA	CGAAAGGGGC	ATTTCTAGGT	6060
GGAGCAGACG	TCAAAGAATG	TATAGAAGCA	TTCAGCAGCT	ATATATGCCC	TTCTGATCCA	6120
GGATTTGTAT	TAAACCATGA	AATAGAGAGC	TGCTTATCAG	GAAACATATC	CCAATGTCCA	6180
AGAACAACGG	TCACATCAGA	CATTGTTCCA	AGATATGCAT	TTGTCAATGG	AGGAGTGGTT	6240
GCAAACCTGTA	TAACAACCAC	CTGTACATGC	AACGGAATTG	GTAATAGAAT	CAATCAACCA	6300
CCTGATCAAG	GAGTAAAAAT	TATAACACAT	AAAGAATGTA	GTACAATAGG	TATCAACGGA	6360
ATGCTGTTCA	ATACAAATAA	AGAAGGAACT	CTTGCATTCT	ATACACCAAA	TGATATAACA	6420
CTAAACAATT	CTGTTGCACT	TGATCCAATT	GACATATCAA	TCGAGCTCAA	CAAGGCCAAA	6480
TCAGATCTAG	AAGAATCAAA	AGAATGGATA	AGAAGGTCAA	ATCAAAAAC	AGATTCTATT	6540
GGAAATTGGC	ATCAATCTAG	CACTACAATC	ATAATTATTT	TGATAATGAT	CATTATATTG	6600
TTTATAATTA	ATATAACGAT	AATTACAATT	GCAATTAAGT	ATTACAGAAT	TCAAAAGAGA	6660
AATCGAGTGG	ATCAAAATGA	CAAGCCATAT	GTACTAACAA	ACAAATAACA	TATCTACAGA	6720
TCATTAGATA	TTAAATTTAT	AAAAAACTTA	GGAGTAAAGT	TACGCAATCC	AACTCTACTC	6780
ATATAATTGA	GGAAGGACCC	AATAGACAAA	TCCAAATTCG	AGATGGAATA	CTGGAAGCAT	6840
ACCAATCACG	GAAAGGATGC	TGGTAATGAG	CTGGAGACGT	CTATGGCTAC	TCATGGCAAC	6900
AAGCTCACTA	ATAAGATAAT	ATACATATTA	TGGACAATAA	TCCTGGTGTT	ATTATCAATA	6960
GTCTTCATCA	TAGTGCTAAT	TAATTCCATC	AAAAGTGAAA	AGGCCACGA	ATCATTGCTG	7020
CAAGACATAA	ATAATGAGTT	TATGGAAATT	ACAGAAAAGA	TCCAAATGGC	ATCGGATAAT	7080
ACCAATGATC	TAATACAGTC	AGGAGTGAAT	ACAAGGCTTC	TTACAATTCA	GAGTCATGTC	7140
CAGAATTACA	TACCAATATC	ATTGACACAA	CAGATGTCAG	ATCTTAGGAA	ATTCATTAGT	7200
GAAATTACAA	TTAGAAATGA	TAATCAAGAA	GTGCTGCCAC	AAAGAATAAC	ACATGATGTA	7260
GGTATAAAAC	CTTTAAATCC	AGATGATTTT	TGGAGATGCA	CGTCTGGTCT	TCCATCTTTA	7320
ATGAAAAC	CAAAAATAAG	GTTAATGCCA	GGGCCGGGAT	TATTAGCTAT	GCCAACGACT	7380
GTTGATGGCT	GTGTTAGAAC	TCCGTCTTTA	GTTATAAATG	ATCTGATTTA	TGCTTATACC	7440
TCAAATCTAA	TTACTCGAGG	TTGTCAGGAT	ATAGGAAAAAT	CATATCAAGT	CTTACAGATA	7500

GGGATAATAA	CTGTAAACTC	AGACTTGGTA	CCTGACTTAA	ATCCTAGGAT	CTCTCATACC	7560
TTTAACATAA	ATGACAATAG	GAAGTCATGT	TCTCTAGCAC	TCCTAAATAT	AGATGTATAT	7620
CAACTGTGTT	CAACTCCCAA	AGTTGATGAA	AGATCAGATT	ATGCATCATC	AGGCATAGAA	7680
GATATTGTAC	TTGATATTGT	CAATTATGAT	GGTTCAATCT	CAACAACAAG	ATTTAAGAAT	7740
AATAACATAA	GCTTTGATCA	ACCATATGCT	GCACTATAACC	CATCTGTTGG	ACCAGGGATA	7800
TACTACAAAG	GCAAAATAAT	ATTTCTCGGG	TATGGAGGTC	TTGAACATCC	AATAAATGAG	7860
AATGTAATCT	GCAACACAAC	TGGGTGCCCC	GGGAAAACAC	AGAGAGACTG	TAATCAAGCA	7920
TCTCATAGTA	CTTGGTTTTTC	AGATAGGAGG	ATGGTCAACT	CCATCATTGT	TGTTGACAAA	7980
GGCTTAAACT	CAATTCCAAA	ATTGAAAGTA	TGGACGATAT	CTATGCGACA	AAATTACTGG	8040
GGGTCAGAAG	GAAGGTTACT	TCTACTAGGT	AACAAGATCT	ATATATATAC	AAGATCTACA	8100
AGTTGGCATA	GCAAGTTACA	ATTAGGAATA	ATTGATATTA	CTGATTACAG	TGATATAAGG	8160
ATAAAATGGA	CATGGCATAA	TGTGCTATCA	AGACCAGGAA	ACAATGAATG	TCCATGGGGA	8220
CATTCATGTC	CAGATGGATG	TATAACAGGA	GTATATACTG	ATGCATATCC	ACTCAATCCC	8280
ACAGGGAGCA	TTGTGTCATC	TGTCATATTA	GACTCACAAA	AATCGAGAGT	GAACCCAGTC	8340
ATAACTTACT	CAACAGCAAC	CGAAAGAGTA	AACGAGCTGG	CCATCCTAAA	CAGAACACTC	8400
TCAGCTGGAT	ATACAACAAC	AAGCTGCATT	ACACACTATA	ACAAAGGATA	TTGTTTTTCAT	8460
ATAGTAGAAA	TAAATCATAA	AAGCTTAAAC	ACATTTCAAC	CCATGTTGTT	CAAAACAGAG	8520
ATTCCAAAAA	GCTGCAGTTA	ATCATAATTA	ACCATAATAT	GCATCAATCT	ATCTATAATA	8580
CAAGTATATG	ATAAGTAATC	AGCAATCAGA	CAATAGACAA	AAGGGAAATA	TAAAAAACTT	8640
AGGAGCAAAG	CGTGCTCGGG	AAATGGACAC	TGAATCTAAC	AATGGCACTG	TATCTGACAT	8700
ACTCTATCCT	GAGTGTCACC	TTAACTCTCC	TATCGTTAAA	GGTAAAATAG	CACAATTACA	8760
CACTATTATG	AGTCTACCTC	AGCCTTATGA	TATGGATGAC	GACTCAATAC	TAGTTATCAC	8820
TAGACAGAAA	ATAAAACTTA	ATAAATTGGA	TAAAAGACAA	CGATCTATTA	GAAGATTAAA	8880
ATTAATATTA	ACTGAAAAAG	TGAATGACTT	AGGAAAATAC	ACATTTATCA	GATATCCAGA	8940
AATGTCAAAA	GAAATGTTCA	AATTATATAT	ACCTGGTATT	AACAGTAAAG	TGACTGAATT	9000
ATTACTTAAA	GCAGATAGAA	CATATAGTCA	AATGACTGAT	GGATTAAGAG	ATCTATGGAT	9060
TAATGTGCTA	TCAAAATTAG	CCTCAAAAAA	TGATGGAAGC	AATTATGATC	TTAATGAAGA	9120
AATTAATAAT	ATATCGAAAAG	TTCACACAAC	CTATAAATCA	GATAAATGGT	ATAATCCATT	9180
CAAAACATGG	TTTACTATCA	AGTATGATAT	GAGAAGATTA	CAAAAAGCTC	GAAATGAGAT	9240
CACTTTTAAT	GTTGGGAAGG	ATTATAACTT	GTTAGAAGAC	CAGAAGAATT	TCTTATTGAT	9300
ACATCCAGAA	TTGGTTTTGA	TATTAGATAA	ACAAAACAT	AATGGTTATC	TAATTACTCC	9360
TGAATTAGTA	TTGATGTATT	GTGACGTAGT	CGAAGGCCGA	TGGAATATAA	GTGCATGTGC	9420
TAAGTTAGAT	CCAAAATTAC	AATCTATGTA	TCAGAAAGGT	AATAACCTGT	GGGAAGTGAT	9480
AGATAAATTG	TTTCCAATTA	TGGGAGAAAA	GACATTTGAT	GTGATATCGT	TATTAGAACC	9540
ACTTGCATTA	TCCTTAATTC	AAACTCATGA	TCCTGTTAAA	CAACTAAGAG	GAGCTTTTTT	9600



AAATCATGTG	TTATCCGAGA	TGGAATTAAT	ATTTGAATCT	AGAGAATCGA	TTAAGGAATT	9660
TCTGAGTGTA	GATTACATTG	ATAAAATTTT	AGATATATTT	AATAAGTCTA	CAATAGATGA	9720
AATAGCAGAG	ATTTTCTCTT	TTTTTAGAAC	ATTTGGGCAT	CCTCCATTAG	AAGCTAGTAT	9780
TGCAGCAGAA	AAGGTTAGAA	AATATATGTA	TATTGGAAAA	CAATTAAAAT	TTGACACTAT	9840
TAATAAATGT	CATGCTATCT	TCTGTACAAT	AATAATTAAC	GGATATAGAG	AGAGGCATGG	9900
TGGACAGTGG	CCTCCTGTGA	CATTACCTGA	TCATGCACAC	GAATTCATCA	TAAATGCTTA	9960
CGGTTCAAAC	TCTGCGATAT	CATATGAAAA	TGCTGTTGAT	TATTACCAGA	GCTTTATAGG	10020
AATAAAATTC	AATAAATTCA	TAGAGCCTCA	GTTAGATGAG	GATTTGACAA	TTTATATGAA	10080
AGATAAAGCA	TTATCTCCAA	AAAAATCAAA	TTGGGACACA	GTTTATCCTG	CATCTAATTT	10140
ACTGTACCGT	ACTAACGCAT	CCAACGAATC	ACGAAGATTA	GTTGAAGTAT	TTATAGCAGA	10200
TAGTAAATTT	GATCCTCATC	AGATATTGGA	TTATGTAGAA	TCTGGGGACT	GGTTAGATGA	10260
TCCAGAATTT	AATATTTCTT	ATAGTCTTAA	AGAAAAAGAG	ATCAAACAGG	AAGGTAGACT	10320
CTTTGCAAAA	ATGACATACA	AAATGAGAGC	TACACAAGTT	TTATCAGAGA	CCCTACTTGC	10380
AAATAACATA	GGAAAATTCT	TTCAAGAAAA	TGGGATGGTG	AAGGGAGAGA	TTGAATTACT	10440
TAAGAGATTA	ACAACCATAT	CAATATCAGG	AGTTCACGG	TATAATGAAG	TGTACAATAA	10500
TTCTAAAAGC	CATACAGATG	ACCTTAAAAC	CTACAATAAA	ATAAGTAATC	TTAATTTGTC	10560
TTCTAATCAG	AAATCAAAGA	AATTTGAATT	CAAGTCAACG	GATATCTACA	ATGATGGATA	10620
CGAGACTGTG	AGCTGTTTCC	TAACAACAGA	TCTCAAAAAA	TACTGTCTTA	ATTGGAGATA	10680
TGAATCAACA	GCTCTATTTG	GAGAAACTTG	CAACCAAATA	TTTGGATTAA	ATAAATTGTT	10740
TAATTGGTTA	CACCCTCGTC	TTGAAGGAAG	TACAATCTAT	GTAGGTGATC	CTTACTGTCC	10800
TCCATCAGAT	AAAGAACATA	TATCATTAGA	GGATCACCCT	GATTCTGGTT	TTTACGTTCA	10860
TAACCCAAGA	GGGGGTATAG	AAGGATTTTG	TCAAAAATTA	TGGACACTCA	TATCTATAAG	10920
TGCAATACAT	CTAGCAGCTG	TTAGAATAGG	CGTGAGGGTG	ACTGCAATGG	TTCAAGGAGA	10980
CAATCAAGCT	ATAGCTGTAA	CCACAAGAGT	ACCCAACAAT	TATGACTACA	GAGTTAAGAA	11040
GGAGATAGTT	TATAAAGATG	TAGTGAGATT	TTTTGATTCA	TTAAGAGAAG	TGATGGATGA	11100
TCTAGGTCAT	GAAC TTAAAT	TAAATGAAAC	GATTATAAGT	AGCAAGATGT	TCATATATAG	11160
CAAAAGAATC	TATTATGATG	GGAGAATTCT	TCCTCAAGCT	CTAAAAGCAT	TATCTAGATG	11220
TGTCTTCTGG	TCAGAGACAG	TAATAGACGA	AACAAGATCA	GCATCTTCAA	ATTTGGCAAC	11280
ATCATTTGCA	AAAGCAATTG	AGAATGGTTA	TTCACCTGTT	CTAGGATATG	CATGCTCAAT	11340
TTTTAAGAAT	ATTCAACAAC	TATATATTGC	CCTTGGGATG	AATATCAATC	CAACTATAAC	11400
ACAGAATATC	AGAGATCAGT	ATTTTAGGAA	TCCAAATTGG	ATGCAATATG	CCTCTTTAAT	11460
ACCTGCTAGT	GTTGGGGGAT	TCAATTACAT	GGCCATGTCA	AGATGTTTTG	TAAGGAATAT	11520
TGGTGATCCA	TCAGTTGCCG	CATTGGCTGA	TATTAAAAGA	TTTATTAAGG	CGAATCTATT	11580
AGACCGAAGT	GTTCTTTATA	GGATTATGAA	TCAAGAACCA	GGTGAGTCAT	CTTTTTTGGG	11640
CTGGGCTTCA	GATCCATATT	CATGCAATTT	ACCACAATCT	CAAAATATAA	CCACCATGAT	11700



TGATAACTGT	AATAAAGATA	ATAAAGGGAA	TAAAATTAAC	AATTTCTGGG	GACTAGCACT	13860
TAAGAACTAT	CAAGTCCTTA	AAATCAGATC	TATAACAAGT	GATTCTGATG	ATAATGATAG	13920
ACTAGATGCT	AATACAAGTG	GTTTGACACT	TCCTCAAGGA	GGGAATTATC	TATCGCATCA	13980
ATTGAGATTA	TTCGGAATCA	ACAGCACTAG	TTGTCTGAAA	GCTCTTGAGT	TATCACAAAT	14040
TTTAATGAAG	GAAGTCAATA	AAGACAAGGA	CAGGCTCTTC	CTGGGAGAAG	GAGCAGGAGC	14100
TATGCTAGCA	TGTTATGATG	CCACATTAGG	ACCTGCAGTT	AATTATTATA	ATTCAGGTTT	14160
GAATATAACA	GATGTAATTG	GTCAACGAGA	ATTGAAAATA	TTTCCTTCAG	AGGTATCATT	14220
AGTAGGTAAA	AAATTAGGAA	ATGTGACACA	GATTCTTAAC	AGGGTAAAAG	TACTGTTCAA	14280
TGGGAATCCT	AATTCAACAT	GGATAGGAAA	TATGGAATGT	GAGAGCTTAA	TATGGAGTGA	14340
ATTAAATGAT	AAGTCCATTG	GATTAGTACA	TTGTGATATG	GAAGGAGCTA	TCGGTAAATC	14400
AGAAGAAACT	GTTCTACATG	AACATTATAG	TGTTATAAGA	ATTACATACT	TGATTGGGGA	14460
TGATGATGTT	GTTTTAGTTT	CCAAAATTAT	ACCTACAATC	ACTCCGAATT	GGTCTAGAAT	14520
ACTTTATCTA	TATAAATTAT	ATTGGAAAGA	TGTAAGTATA	ATATCACTCA	AAACTTCTAA	14580
TCCTGCATCA	ACAGAATTAT	ATCTAATTTT	GAAAGATGCA	TATTGTACTA	TAATGGAACC	14640
TAGTGAAATT	GTTTTATCAA	AACTTAAAAG	ATTGTCACTC	TTGGAAGAAA	ATAATCTATT	14700
AAAATGGATC	ATTTTATCAA	AGAAGAGGAA	TAATGAATGG	TTACATCATG	AAATCAAAGA	14760
AGGAGAAAGA	GATTATGGAA	TCATGAGACC	ATATCATATG	GCACTACAAA	TCTTTGGATT	14820
TCAAATCAAT	TTAAATCATC	TGGCGAAAGA	ATTTTTATCA	ACCCAGATC	TGACTAATAT	14880
CAACAATATA	ATCCAAAGTT	TTCAGCGAAC	AATAAAGGAT	GTTTTATTTG	AATGGATTAA	14940
TATAACTCAT	GATGATAAGA	GACATAAATT	AGGCGGAAGA	TATAACATAT	TCCCACTGAA	15000
AAATAAGGGA	AAGTTAAGAC	TGCTATCGAG	AAGACTAGTA	TTAAGTTGGA	TTTCATTATC	15060
ATTATCGACT	CGATTACTTA	CAGGTCGCTT	TCCTGATGAA	AAATTTGAAC	ATAGAGCACA	15120
GACTGGATAT	GTATCATTAG	CTGATACTGA	TTTAGAATCA	TTAAAGTTAT	TGTCGAAAAA	15180
CATCATTAA	AATTACAGAG	AGTGTATAGG	ATCAATATCA	TATTGGTTTC	TAACCAAAGA	15240
AGTTAAAATA	CTTATGAAAT	TGATCGGTGG	TGCTAAATTA	TTAGGAATTC	CCAGACAATA	15300
TAAAGAACCC	GAAGACCAGT	TATTAGAAAA	CTACAATCAA	CATGATGAAT	TTGATATCGA	15360
TTAAAACATA	AATACAATGA	AGATATATCC	TAACCTTTAT	CTTTAAGCCT	AGGAATAGAC	15420
AAAAAGTAAG	AAAAACATGT	AATATATATA	TACCAAACAG	AGTTCTTCTC	TTGTTTGGTG	15480
GGTCGGCATG	GCATCTCCAC	CTCCTCGCGG	TCCGGACCTG	GGCATCCGAA	GGAGGACGCA	15540
CGTCCACTCG	GATGGCTAAG	GGAGAGCCTG	CAGTAGCATA	ACCCCTTGGG	GCCTCTAAAC	15600
GGGTCTTGAG	GGGTTTTTTT	CTGAAAGGAG	GAAGTATATA	CGCGTCGACG	GGCCCCGCGC	15660

## (2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 15666 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TAATACGACT	CACTATAGGA	CCAAACAAGA	GAAGAACTT	GTCTGGAAT	ATAAATTAA	60
CTTTAAATTA	ACTTAGGATT	AAAGACATTG	ACTAGAAGGT	CAAGAAAAGG	GAAGCTCTATA	120
ATTTCAAAAA	TGTTGAGCCT	ATTTGATACA	TTTAATGCAC	GTAGGCAAGA	AAACATAACA	180
AAATCAGCCG	GTGGAGCTAT	CATTCTGGA	CAGAAAATA	CTGTCTCTAT	ATTCGCCCTT	240
GGACCGACAA	TAAGTATGA	TAATGAGAAA	ATGACATTAG	CTCTTCTATT	TCTATCTCAT	300
TCACTAGATA	ATGAGAAACA	ACATGCACAA	AGGGCAGGGT	TCTTGGTGTC	TTTATTGTCA	360
ATGGCTTATG	CCAATCCAGA	GCTCTACCTA	ACAACAAATG	GAAGTAATGC	AGATGTCAAG	420
TATGTCATAT	ACATGATTGA	GAAAGATCTA	AAACGGCAAA	AGTATGGAGG	ATTTGTGGTT	480
AAGACGAGAG	AGATGATATA	TGAAAAGACA	ACTGATTGGA	TATTTGGAAG	TGACCTGGAT	540
TATGATCAGG	AAACTATGTT	GCAGAACGGC	AGGAACAATT	CAACAATTGA	AGACCTTGTC	600
CACACATTTG	GGTATCCATC	ATGTTTAGGA	GCTCTTATAA	TACAGATCTG	GATAGTTCTG	660
GTCAAAGCTA	TCACTAGTAT	CTCAGGGTTA	AGAAAAGGCT	TTTTCAACCG	ATTGGAAGCT	720
TTCAGACAAG	ATGGAACAGT	GCAGGCAGGG	CTGGTATTGA	GCGGTGACAC	AGTGGATCAG	780
ATTGGGTCAA	TCATGCGGTC	TCAACAGAGC	TTGGTAACTC	TTATGGTTGA	AACATTAATA	840
ACAATGAATA	CCAGCAGAAA	TGACCTCACA	ACCATAGAAA	AGAATATACA	AATTGTTGGC	900
AACTACATAA	GAGATGCAGG	TCTCGCTTCA	TTCTTCAATA	CAATCAGATA	TGGAATTGAG	960
ACCAGAATGG	CAGCTTTGAC	TCTATCCACT	CTCAGACCAG	ATATCAATAG	ATTAAAAGCT	1020
TTGATGGAAC	TGTATTTATC	AAAGGGACCA	CGCGCTCCTT	TCATCTGTAT	CCTCAGAGAT	1080
CCTATACATG	GTGAGTTTCG	ACCAGGCAAC	TATCCTGCCA	TATGGAGCTA	TGCAATGGGG	1140
GTGGCAGTTG	TACAAAATAG	AGCCATGCAA	CAGTATGTGA	CGGGAAGATC	ATATCTAGAC	1200
ATTGATATGT	TCCAGCTAGG	ACAAGCAGTA	GCACGTGATG	CCGAAGCTCA	AATGAGCTCA	1260
ACACTGGAAG	ATGAAGTTGG	AGTGACACAC	GAATCTAAAG	AAAGCTTGAA	GAGACATATA	1320
AGGAACATAA	ACAGTTCAGA	GACATCTTTC	CACAAACCGA	CAGGTGGATC	AGCCATAGAG	1380
ATGGCAATAG	ATGAAGAGCC	AGAACAATTC	GAACATAGAG	CAGATCAAGA	ACAAAATGGA	1440
GAACCTCAAT	CATCCATAAT	TCAATATGCC	TGGGCAGAAG	GAAATAGAAG	CGATGATCAG	1500
ACTGAGCAAG	CTACAGAATC	TGACAATATC	AAGACCGAAC	AACAAAACAT	CAGAGACAGA	1560
CTAAACAAGA	GACTCAACGA	CAAGAAGAAA	CAAAGCAGTC	AACCACCCAC	TAATCCCACA	1620
AACAGAACAA	ACCAGGACGA	AATAGATGAT	CTGTTTAACG	CATTTGGAAG	CAACTAATCG	1680
AATCAACATT	TTAATCTAAA	TCAATAATAA	ATAAGAAAAA	CTTAGGATTA	AAGAATCCTA	1740
TCATACCGGA	ATATAGGGTG	GTAAATTTAG	AGTCTGCTTG	AAACTCAATC	AATAGAGAGT	1800
TGATGGAAAG	CGATGCTAAA	AACTATCAAA	TCATGGATTTC	TTGGGAAGAG	GAATCAAGAG	1860

ATAAATCAAC	TAATATCTCC	TCGGCCCTCA	ACATCATTTGA	ATTCTACTCT	AGCACCGACC	1920
CCCAAGAAGA	CTTATCGGAA	AACGACACAA	TCAACACAAG	AACCCAGCAA	CTCAGTGCCA	1980
CCATCTGTCA	ACCAGAAATC	AAACCAACAG	AAACAAGTGA	GAAAGATAGT	GGATCAACTG	2040
ACAAAAATAG	ACAGTCTGGG	TCATCACACG	AATGTACAAC	AGAAGCAAAA	GATAGAAATA	2100
TTGATCAGGA	AACTGTACAG	AGAGGACCTG	GGAGAAGAAG	CAGCTCAGAT	AGTAGAGCTG	2160
AGACTGTGGT	CTCTGGAGGA	ATCCCCAGAA	GCATCACAGA	TTCTAAAAAT	GGAACCCAAA	2220
ACACGGAGGA	TATTGATCTC	AATGAAATTA	GAAAGATGGA	TAAGGACTCT	ATTGAGGGGA	2280
AAATGCGACA	ATCTGCAAAT	GTTCCAAGCG	AGATATCAGG	AAGTGATGAC	ATATTTACAA	2340
CAGAACAAAG	TAGAAACAGT	GATCATGGAA	GAAGCCTGGA	ATCTATCAGT	ACACCTGATA	2400
CAAGATCAAT	AAGTGTTGTT	ACTGCTGCAA	CACCAGATGA	TGAAGAAGAA	ATACTAATGA	2460
AAAATAGTAG	GACAAAGAAA	AGTTCTTCAA	CACATCAAGA	AGATGACAAA	AGAATTAAAA	2520
AAGGGGGAAA	AGGGAAAGAC	TGGTTTAAGA	AATCAAAAAG	TACCGACAAC	CAGATACCAA	2580
CATCAGACTA	CAGATCCACA	TCAAAAGGGC	AGAAGAAAAT	CTCAAAGACA	ACAACCACCA	2640
ACACCGACAC	AAAGGGGCAA	ACAGAAATAC	AGACAGAATC	ATCAGAAACA	CAATCCTCAT	2700
CATGGAATCT	CATCATCGAC	AACAACACCG	ACCGGAACGA	ACAGACAAGC	ACAACCTCTC	2760
CAACAACAAC	TTCCAGATCA	ACTTATACAA	AAGAATCGAT	CCGAACAAAC	TCTGAATCCA	2820
AACCCAAGAC	ACAAAAGACA	AATGGAAAAG	AAAGGAAGGA	TACAGAAGAG	AGCAATCGAT	2880
TTACAGAGAG	GGCAATTACT	CTATTGCAGA	ATCTTGGTGT	AATTCAATCC	ACATCAAAAC	2940
TAGATTTATA	TCAAGACAAA	CGAGTTGTAT	GTGTAGCAAA	TGTACTAAAC	AATGTAGATA	3000
CTGCATCAAA	GATAGATTTT	CTGGCAGGAT	TAGTCATAGG	GGTTTCAATG	GACAACGACA	3060
CAAAATTAAC	ACAGATACAA	AATGAAATGC	TAAACCTCAA	AGCAGATCTA	AAGAAAATGG	3120
ACGAATCACA	TAGAAGATTG	ATAGAAAATC	AAAGAGAACA	ACTGTCATTG	ATCACGTCAC	3180
TAATTTCAAA	TCTCAAAATT	ATGACTGAGA	GAGGAGGAAA	GAAAGACCAA	AATGAATCCA	3240
ATGAGAGAGT	ATCCATGATC	AAAACAAAAT	TGAAAGAAGA	AAAGATCAAG	AAGACCAGGT	3300
TTGACCCACT	TATGGAGGCA	CAAGGCATTG	ACAAGAATAT	ACCCGATCTA	TATCGACATG	3360
CAGGAGATAC	ACTAGAGAAC	GATGTACAAG	TTAAATCAGA	GATATTAAGT	TCATACAATG	3420
AGTCAAATGC	AACAAGACTA	ATACCCAAAA	AAGTGAGCAG	TACAATGAGA	TCACTAGTTG	3480
CAGTCATCAA	CAACAGCAAT	CTCTCACAAA	GCACAAAACA	ATCATACATA	AACGAACTCA	3540
AACGTTGCAA	AAATGATGAA	GAAGTATCTG	AATTAATGGA	CATGTTCAAT	GAAGATGTCA	3600
ACAATTGCCA	ATGATCCAAC	AAAGAAACGA	CACCGAACAA	ACAGACAAGA	AACAACAGTA	3660
GATCAAAACC	TGTCAACACA	CACAAAATCA	AGCAGAATGA	AACAACAGAT	ATCAATCAAT	3720
ATACAAATAA	GAAAAACTTA	GGATTAAAGA	ATAAATTAAT	CCTTGTCCAA	AATGAGTATA	3780
ACTAACTCTG	CAATATACAC	ATTCCCAGAA	TCATCATTTCT	CTGAAAATGG	TCATATAGAA	3840
CCATTACCAC	TCAAAGTCAA	TGAACAGAGG	AAAGCAGTAC	CCCACATTAG	AGTTGCCAAG	3900
ATCGGAAATC	CACCAAAAACA	CGGATCCCGG	TATTTAGATG	TCTTCTTACT	CGGCTTCTTC	3960

GAGATGGAAC	GAATCAAAGA	CAAATACGGG	AGTGTGAATG	ATCTCGACAG	TGACCCCGAGT	4020
TACAAAGTTT	GTGGCTCTGG	ATCATTACCA	ATCGGATTGG	CTAAGTACAC	TGGGAATGAC	4080
CAGGAATTGT	TACAAGCCGC	AACCAAACCTG	GATATAGAAG	TGAGAAGAAC	AGTCAAAGCG	4140
AAAGAGATGG	TTGTTTACAC	GGTACAAAAT	ATAAAACCAG	AACTGTACCC	ATGGTCCAAT	4200
AGACTAAGAA	AAGGAATGCT	GTTGATGCC	AACAAAGTTG	CTCTTGCTCC	TCAATGTCTT	4260
CCACTAGATA	GGAGCATAAA	ATTTAGAGTA	ATCTTCGTGA	ATTGTACGGC	AATTGGATCA	4320
ATAACCTTGT	TCAAAATTCC	TAAGTCAATG	GCATCACTAT	CTCTACCCAA	CACAATATCA	4380
ATCAATCTGC	AGGTACACAT	AAAAACAGGG	GTTCAGACTG	ATTCTAAAGG	GATAGTTCAA	4440
ATTTTGGATG	AGAAAGGCGA	AAAATCACTG	AATTTCATGG	TCCATCTCGG	ATTGATCAAA	4500
AGAAAAGTAG	GCAGAATGTA	CTCTGTTGAA	TACTGTAAAC	AGAAAATCGA	GAAAATGAGA	4560
TTGATATTTT	CTTTAGGACT	AGTTGGAGGA	ATCAGTCTTC	ATGTCAATGC	AACTGGGTCC	4620
ATATCAAAAA	CACTAGCAAG	TCAGCTGGTA	TTCAAAAGAG	AGATTTGTTA	TCCTTTAATG	4680
GATCTAAATC	CGCATCTCAA	TCTAGTTATC	TGGGCTTCAT	CAGTAGAGAT	TACAAGAGTG	4740
GATGCAATTT	TCCAACCTTC	TTTACCTGGC	GAGTTCAGAT	ACTATCCTAA	TATTATTGCA	4800
AAAGGAGTTG	GGAAAATCAA	ACAATGGAAC	TAGTAATCTC	TATTTTAGTC	CGGACGTATC	4860
TATTAAGCCG	AAGCAAATAA	AGGATAATCA	AAAACCTAGG	ACAAAAGAGG	TCAATACCAA	4920
CAACTATTAG	CAGTCACACT	CGCAAGAATA	AGAGAGAAGG	GACCAAAAAA	GTCAAATAGG	4980
AGAAATCAAA	ACAAAAGGTA	CAGAACACCA	GAACAACAAA	ATCAAAACAT	CCAACTCACT	5040
CAAAACAAAA	ATTCCAAAAG	AGACCGGCAA	CACAACAAGC	ACTGAACACA	ATGCCAACTT	5100
CAATACTGCT	AATTATTACA	ACCATGATCA	TGGCATCTTT	CTGCCAAATA	GATATCACAA	5160
AACTACAGCA	CGTAGGTGTA	TTGGTCAACA	GTCCCAAAGG	GATGAAGATA	TCACAAAACCT	5220
TTGAAACAAG	ATATCTAATT	TTGAGCCTCA	TACCAAAAAT	AGAAGACTCT	AACTCTTG TG	5280
GTGACCAACA	GATCAAGCAA	TACAAGAAGT	TATTGGATAG	ACTGATCATC	CCTTTATATG	5340
ATGGATTAAG	ATTACAGAAA	GATGTGATAG	TAACCAATCA	AGAATCCAAT	GAAAACACTG	5400
ATCCCAGAAC	AAAACGATTC	TTTGGAGGGG	TAATTGGAAC	CATTGCTCTG	GGAGTAGCAA	5460
CCTCAGCACA	AATTACAGCG	GCAGTTGCTC	TGGTTGAAGC	CAAGCAGGCA	AGATCAGACA	5520
TCGAAAAACT	CAAAGAAGCA	ATTAGGGACA	CAAACAAAGC	AGTGCAGTCA	G TTCAGAGCT	5580
CCATAGGAAA	TTTAATAGTA	GCAATTAAAT	CAGTCCAGGA	TTATGTTAAC	AAAGAAATCG	5640
TGCCATCGAT	TGCGAGGCTA	GGTTGTGAAG	CAGCAGGACT	TCAATTAGGA	ATTGCATTAA	5700
CACAGCATT A	CTCAGAATTA	ACAAACATAT	TTGGTGATAA	CATAGGATCG	TTACAAGAAA	5760
AAGGAATAAA	ATTACAAGGT	ATAGCATCAT	TATACCGCAC	AAATATCACA	GAAATATTCA	5820
CAACATCAAC	AGTTGATAAA	TATGATATCT	ATGATCTGTT	ATTTACAGAA	TCAATAAAGG	5880
TGAGAGTTAT	AGATGTTGAC	TTGAATGATT	ACTCAATCAC	CCTCCAAGTC	AGACTCCCTT	5940
TATTAAC TAG	GCTGCTGAAC	ACTCAGATCT	ACAAAGTAGA	TTCCATATCA	TATAACATCC	6000
AAAACAGAGA	ATGGTATATC	CCTCTTCCCA	GCCATATCAT	GACGAAAAGG	GCATTTCTAG	6060

GTGGAGCAGA	CGTCAAAGAA	TGTATAGAAG	CATTTCAGCAG	CTATATATGC	CCTTCTGATC	6120
CAGGATTTGT	ATTAAACCAT	GAAATAGAGA	GCTGCTTATC	AGGAAACATA	TCCCAATGTC	6180
CAAGAACAAC	GGTCACATCA	GACATTG'TTC	CAAGATATGC	ATTTGTCAAT	GGAGGAGTGG	6240
TTGCAAAC TG	TATAACAACC	ACCTGTACAT	GCAACGGAAT	TGGTAATAGA	ATCAATCAAC	6300
CACCTGATCA	AGGAGTAAAA	ATTATAACAC	ATAAAGAATG	TAGTACAATA	GGTATCAACG	6360
GAATGCTGTT	CAATACAAAT	AAAGAAGGAA	CTCTTGCATT	CTATACACCA	AATGATATAA	6420
CACTAAACAA	TTCTGTTGCA	CTTGATCCAA	TTGACATATC	AATCGAGCTC	AACAAGGCCA	6480
AATCAGATCT	AGAAGAATCA	AAAGAATGGA	TAAGAAGGTC	AAATCAAAAA	CTAGATTCTA	6540
TTGGAAATTG	GCATCAATCT	AGCACTACAA	TCATAATTAT	TTTGATAATG	ATCATTATAT	6600
TG'TTTATAAT	TAATATAACG	ATAATTACAA	TTGCAATTAA	GTATTACAGA	ATTCAAAAAGA	6660
GAAATCGAGT	GGATCAAAAT	GACAAGCCAT	ATGTACTAAC	AAACAAATAA	CATATCTACA	6720
GATCATTAGA	TATTAAAATT	ATAAAAACT	TAGGAGTAAA	GTTACGCAAT	CCAACTCTAC	6780
TCATATAATT	GAGGAAGGAC	CCAATAGACA	AATCCAAATT	CGAGATGGAA	TACTGGAAGC	6840
ATACCAATCA	CGGAAAGGAT	GCTGGTAATG	AGCTGGAGAC	GTCTATGGCT	ACTCATGGCA	6900
ACAAGCTCAC	TAATAAGATA	ATATACATAT	TATGGACAAT	AATCCTGGTG	TTATTATCAA	6960
TAGTCTTCAT	CATAGTGCTA	ATTAATTCCA	TCAAAGTGA	AAAGGCCAC	GAATCATTGC	7020
TGCAAGACAT	AAATAATGAG	TTTATGGAAA	TTACAGAAAA	GATCCAAATG	GCATCGGATA	7080
ATACCAATGA	TCTAATACAG	TCAGGAGTGA	ATACAAGGCT	TCTTACAATT	CAGAGTCATG	7140
TCCAGAATTA	CATACCAATA	TCATTGACAC	AACAGATGTC	AGATCTTAGG	AAATTCATTA	7200
GTGAAATTAC	AATTAGAAAT	GATAATCAAG	AAGTGCTGCC	ACAAAGAATA	ACACATGATG	7260
TAGGTATAAA	ACCTTTAAAT	CCAGATGATT	TTTGGAGATG	CACGTCTGGT	CTTCCATCTT	7320
TAATGAAAAC	TCCAAAAATA	AGGTTAATGC	CAGGGCCGGG	ATTATTAGCT	ATGCCAACGA	7380
CTGTTGATGG	CTGTGTTAGA	ACTCCGTCTT	TAGTTATAAA	TGATCTGATT	TATGCTTATA	7440
CCTCAAATCT	AATTACTCGA	GGTTGTCAGG	ATATAGGAAA	ATCATATCAA	GTCTTACAGA	7500
TAGGGATAAT	AACTGTAAAC	TCAGACTTGG	TACCTGACTT	AAATCCTAGG	ATCTCTCATA	7560
CCTTTAACAT	AAATGACAAT	AGGAAGTCAT	GTTCTCTAGC	ACTCCTAAAT	ATAGATGTAT	7620
ATCAACTGTG	TTCAACTCCC	AAAGTTGATG	AAAGATCAGA	TTATGCATCA	TCAGGCATAG	7680
AAGATATTGT	ACTTGATATT	GTCAATTATG	ATGGTTCAAT	CTCAACAACA	AGATTTAAGA	7740
ATAATAACAT	AAGCTTTGAT	CAACCATATG	CTGCACTATA	CCCATCTGTT	GGACCAGGGA	7800
TATACTACAA	AGGCAAAATA	ATATTTCTCG	GGTATGGAGG	TCTTGAACAT	CCAATAAATG	7860
AGAATGTAAT	CTGCAACACA	ACTGGGTGCC	CCGGGAAAAC	ACAGAGAGAC	TGTAATCAAG	7920
CATCTCATAG	TACTTGGTTT	TCAGATAGGA	GGATGGTCAA	CTCCATCATT	GTTGTTGACA	7980
AAGGCTTAAA	CTCAATTCCA	AAATTGAAAG	TATGGACGAT	ATCTATGCGA	CAAAATTACT	8040
GGGGGT CAGA	AGGAAGGTTA	CTTCTACTAG	GTAACAAGAT	CTATATATAT	ACAAGATCTA	8100
CAAGTTGGCA	TAGCAAGTTA	CAATTAGGAA	TAATTGATAT	TACTGATTAC	AGTGATATAA	8160

GGATAAAATG	GACATGGCAT	AATGTGCTAT	CAAGACCAGG	AAACAATGAA	TGTCCATGGG	8220
GACATTCATG	TCCAGATGGA	TGTATAACAG	GAGTATATAC	TGATGCATAT	CCACTCAATC	8280
CCACAGGGAG	CATTGTGTCA	TCTGTCATAT	TAGACTCACA	AAAATCGAGA	GTGAACCCAG	8340
TCATAACTTA	CTCAACAGCA	ACCGAAAGAG	TAAACGAGCT	GGCCATCCTA	AACAGAACAC	8400
TCTCAGCTGG	ATATACAACA	ACAAGCTGCA	TTACACACTA	TAACAAAGGA	TATTGTTTTC	8460
ATATAGTAGA	AATAAATCAT	AAAAGCTTAA	ACACATTTCA	ACCCATGTTG	TTCAAAACAG	8520
AGATTCCAAA	AAGCTGCAGT	TAATCATAAT	TAACCATAAT	ATGCATCAAT	CTATCTATAA	8580
TACAAGTATA	TGATAAGTAA	TCAGCAATCA	GACAATAGAC	AAAAGGGAAA	TATAAAAAAC	8640
TTAGGAGCAA	AGCGTGCTCG	GGAAATGGAC	ACTGAATCTA	ACAATGGCAC	TGTATCTGAC	8700
ATACTCTATC	CTGAGTGTCA	CCTTAAC'TCT	CCTATCGTTA	AAGGTAAAAT	AGCACAATTA	8760
CACACTATTA	TGAGTCTACC	TCAGCCTTAT	GATATGGATG	ACGACTCAAT	ACTAGTTATC	8820
ACTAGACAGA	AAATAAAACT	TAATAAATTG	GATAAAAGAC	AACGATCTAT	TAGAAGATTA	8880
AAATTAATAT	TAAGTAAAA	AGTGAATGAC	TTAGGAAAAT	ACACATTTAT	CAGATATCCA	8940
GAAATGTCAA	AAGAAATGTT	CAAATTATAT	ATACCTGGTA	TTAACAGTAA	AGTGAAGTAA	9000
TTATTACTTA	AAGCAGATAG	AACATATAGT	CAAATGACTG	ATGGATTAAG	AGATCTATGG	9060
ATTAATGTGC	TATCAAAAATT	AGCCTCAAAA	AATGATGGAA	GCAATTATGA	TCTTAATGAA	9120
GAAATTAATA	ATATATCGAA	AGTTCACACA	ACCTATAAAT	CAGATAAATG	GTATAATCCA	9180
TTCAAAACAT	GGTTTACTAT	CAAGTATGAT	ATGAGAAGAT	TACAAAAAGC	TCGAAATGAG	9240
ATCACTTTTA	ATGTTGGGAA	GGATTATAAC	TTGTTAGAAG	ACCAGAAGAA	TTTCTTATTG	9300
ATACATCCAG	AATTGGTTTT	GATATTAGAT	AAACAAAAC	ATAATGGTTA	TCTAATTACT	9360
CCTGAATTAG	TATTGATGTA	TTGTGACGTA	GTCGAAGGCC	GATGGAATAT	AAGTGCATGT	9420
GCTAAGTTAG	ATCCAAAATT	ACAATCTATG	TATCAGAAAG	GTAATAACCT	GTGGGAAGTG	9480
ATAGATAAAT	TGTTTCCAAT	TATGGGAGAA	AAGACATTTG	ATGTGATATC	GTTATTAGAA	9540
CCACTTGCAT	TATCCTTAAT	TCAAACTCAT	GATCCTGTTA	AACAACCTAG	AGGAGCTTTT	9600
TTAAATCATG	TGTTATCCGA	GATGGAATTA	ATATTTGAAT	CTAGAGAATC	GATTAAGGAA	9660
TTTCTGAGTG	TAGATTACAT	TGATAAAATT	TTAGATATAT	TTAATAAGTC	TACAATAGAT	9720
GAAATAGCAG	AGATTTTCTC	TTTTTTT'TAGA	ACATTTGGGC	ATCCTCCATT	AGAAGCTAGT	9780
ATTGCAGCAG	AAAAGGTTAG	AAAATATATG	TATATTGGAA	AACAATTAAA	ATTTGACACT	9840
ATTAATAAAT	GTCATGCTAT	CTTCTGTACA	ATAATAATTA	ACGGATATAG	AGAGAGGCAT	9900
GGTGGACAGT	GGCCTCCTGT	GACATTACCT	GATCATGCAC	ACGAATTCAT	CATAAATGCT	9960
TACGGTTCAA	ACTCTGCGAT	ATCATATGAA	AATGCTGTTG	ATTATTACCA	GAGCTTTATA	10020
GGAATAAAAT	TCAATAAATT	CATAGAGCCT	CAGTTAGATG	AGGATTTGAC	AATTTATATG	10080
AAAGATAAAG	CATTATCTCC	AAAAAAATCA	AATTGGGACA	CAGTTTATCC	TGCATCTAAT	10140
TTACTGTACC	GTACTAACGC	ATCCAACGAA	TCACGAAGAT	TAGTTGAAGT	ATTTATAGCA	10200
GATAGTAAAT	TTGATCCTCA	TCAGATATTG	GATTATGTAG	AATCTGGGGA	CTGGTTAGAT	10260



GATCCAGAAT	TTAATATTTT	TTATAGTCTT	AAAGAAAAAG	AGATCAAACA	GGAAGGTAGA	10320
CTCTTTGCAA	AAATGACATA	CAAAATGAGA	GCTACACAAG	TTTTATCAGA	GACCCTACTT	10380
GCAAATAACA	TAGGAAAATT	CTTTCAAGAA	AATGGGATGG	TGAAGGGAGA	GATTGAATTA	10440
CTTAAGAGAT	TAACAACCAT	ATCAATATCA	GGAGTTCCAC	GGTATAATGA	AGTGTACAAAT	10500
AATTCTAAAA	GCCATACAGA	TGACCTTAAA	ACCTACAATA	AAATAAGTAA	TCTTAATTTG	10560
TCTTCTAATC	AGAAATCAAA	GAAATTTGAA	TTCAAGTCAA	CGGATATCTA	CAATGATGGA	10620
TACGAGACTG	TGAGCTGTTT	CCTAACAACA	GATCTCAAAA	AATACTGTCT	TAATTGGAGA	10680
TATGAATCAA	CAGCTCTATT	TGGAGAAACT	TGCAACCAAA	TATTTGGATT	AAATAAATTG	10740
TTTAATTGGT	TACACCCTCG	TCTTGAAGGA	AGTACAATCT	ATGTAGGTGA	TCCTTACTGT	10800
CCTCCATCAG	ATAAAGAACA	TATATCATT	GAGGATCACC	CTGATTCTGG	TTTTTACGTT	10860
CATAACCCAA	GAGGGGGTAT	AGAAGGATTT	TGTCAAAAAT	TATGGACACT	CATATCTATA	10920
AGTGCAATAC	ATCTAGCAGC	TGTTAGAATA	GGCGTGAGGG	TGACTGCAAT	GGTTCAAGGA	10980
GACAATCAAG	CTATAGCTGT	AACCACAAGA	GTACCCAACA	ATTATGACTA	CAGAGTTAAG	11040
AAGGAGATAG	TTTATAAAGA	TGTAGTGAGA	TTTTTTTGATT	CATTAAGAGA	AGTGATGGAT	11100
GATCTAGGTC	ATGAACTTAA	ATTAAATGAA	ACGATTATAA	GTAGCAAGAT	GTTTCATATAT	11160
AGCAAAAGAA	TCTATTATGA	TGGGAGAATT	CTTCCTCAAG	CTCTAAAAGC	ATTATCTAGA	11220
TGTGTCTTCT	GGTCAGAGAC	AGTAATAGAC	GAAACAAGAT	CAGCATCTTC	AAATTTGGCA	11280
ACATCATTTG	CAAAAGCAAT	TGAGAATGGT	TATTCACCTG	TTCTAGGATA	TGCATGCTCA	11340
ATTTTTAAGA	ATATTCAACA	ACTATATATT	GCCCTTGGA	TGAATATCAA	TCCAACCTATA	11400
ACACAGAATA	TCAGAGATCA	GTATTTTAGG	AATCCAAATT	GGATGCAATA	TGCCTCTTTA	11460
ATACCTGCTA	GTGTTGGGGG	ATTCAATTAC	ATGGCCATGT	CAAGATGTTT	TGTAAGGAAT	11520
ATTGGTGATC	CATCAGTTGC	CGCATTGGCT	GATATTAAAA	GATTTATTAA	GGCGAATCTA	11580
TTAGACCGAA	GTGTTCTTTA	TAGGATTATG	AATCAAGAAC	CAGGTGAGTC	ATCTTTTTTTG	11640
GACTGGGCTT	CAGATCCATA	TTCATGCAAT	TTACCACAAT	CTCAAAATAT	AACCACCATG	11700
ATAAAAAATA	TAACAGCAAG	GAATGTATTA	CAAGATTCAC	CAAATCCATT	ATTATCTGGA	11760
TTATTCACAA	ATACAATGAT	AGAAGAAGAT	GAAGAATTAG	CTGAGTTCCT	GATGGACAGG	11820
AAGGTAATTC	TCCCTAGAGT	TGCACATGAT	ATTCTAGATA	ATTCTCTCAC	AGGAATTAGA	11880
AATGCCATAG	CTGGAATGTT	AGATACGACA	AAATCACTAA	TTCGGGTTGG	CATAAATAGA	11940
GGAGGACTGA	CATATAGTTT	GTTGAGGAAA	ATCAGTAATT	ACGATCTAGT	ACAATATGAA	12000
ACACTAAGTA	GGACTTTGCG	ACTAATTGTA	AGTGATAAAA	TCAAGTATGA	AGATATGTGT	12060
TCGGTAGACC	TTGCCATAGC	ATTGCGACAA	AAGATGTGGA	TTCATTTATC	AGGAGGAAGG	12120
ATGATAAGTG	GACTTGAAAC	GCCTGACCCA	TTAGAATTAC	TATCTGGGGT	AGTAATAACA	12180
GGATCAGAAC	ATTGTAAAAAT	ATGTTATTCT	TCAGATGGCA	CAAACCCATA	TACTTGATG	12240
TATTTACCCG	GTAATATCAA	AATAGGATCA	GCAGAAACAG	GTATATCGTC	ATTAAGAGTT	12300
CCTTATTTTG	GATCAGTCAC	TGATGAAAGA	TCTGAAGCAC	AATTAGGATA	TATCAAGAAT	12360

CTTAGTAAAC	CTGCAAAAGC	CGCAATAAGA	ATAGCAATGA	TATATACATG	GGCATTGTTG	12420
AATGATGAGA	TATCTTGGAT	GGAAGCCTCA	CAGATAGCAC	AAACACGTGC	AAATTTTACA	12480
CTAGATAGTC	TCAAAATTTT	AACACCGGTA	GCTACATCAA	CAAATTTATC	ACACAGATTA	12540
AAGGATACTG	CAACTCAGAT	GAAATTCTCC	AGTACATCAT	TGATCAGAGT	CAGCAGATTC	12600
ATAACAATGT	CCAATGATAA	CATGTCTATC	AAAGAAGCTA	ATGAAACCAA	AGATACTAAT	12660
CTTATTTATC	AACAAATAAT	GTTAACAGGA	TTAAGTGTTT	TCGAATATTT	ATTTAGATTA	12720
AAAGAAACCA	CAGGACACAA	CCCTATAGTT	ATGCATCTGC	ACATAGAAGA	TGAGTGTTGT	12780
ATTAAAGAAA	GTTTTAATGA	TGAACATATT	AATCCAGAGT	CTACATTAGA	ATTAATTCGA	12840
TATCCTGAAA	GTAATGAATT	TATTTATGAT	AAAGACCCAC	TCAAAGATGT	GGACTTATCA	12900
AAACTTATGG	TTATTAAAGA	CCATTCTTAC	ACAATTGATA	TGAATTATTG	GGATGATACT	12960
GACATCATAC	ATGCAATTTT	AATATGTACT	GCAATTACAA	TAGCAGATAC	TATGTCACAA	13020
TTAGATCGAG	ATAATTTAAA	AGAGATAATA	GTTATTGCAA	ATGATGATGA	TATTAATAGC	13080
TTAATCACTG	AATTTTGTAC	TCTTGACATA	CTTGTATTTT	TCAAGACATT	TGGTGGATTA	13140
TTAGTAAATC	AATTTGCATA	CACTCTTTAT	AGTCTAAAAA	TAGAAGGTAG	GGATCTCATT	13200
TGGGATTATA	TAATGAGAAC	ACTGAGAGAT	ACTTCCCATT	CAATATTAAA	AGTATTATCT	13260
AATGCATTAT	CTCATCCTAA	AGTATTCAAG	AGGTTCTGGG	ATTGTGGAGT	TTTAAACCCT	13320
ATTTATGGTC	CTAATACTGC	TAGTCAAGAC	CAGATAAAAC	TTGCCCTATC	TATATGTGAA	13380
TATTTACTAG	ATCTATTTAT	GAGAGAATGG	TTGAATGGTG	TATCACTTGA	AATATACATT	13440
TGTGACAGCG	ATATGGAAGT	TGCAAATGAT	AGGAAACAAG	CCTTTATTTT	TAGACACCTT	13500
TCATTTGTTT	GTTGTTTAGC	AGAAATTGCA	TCTTTCGGAC	CTAACCTGTT	AAACTTAACA	13560
TACTTGGAGA	GACTTGATCT	ATTGAAACAA	TATCTTGAAT	TAAATATTAA	AGAAGACCCT	13620
ACTCTTAAAT	ATGTACAAAT	ATCTGGATTA	TTAATTAAAT	CGTTCCCATC	AACTGTAACA	13680
TACGTAAGAA	AGACTGCAAT	CAAATATCTA	AGGATTGCGG	GTATTAGTCC	ACCTGAGGTA	13740
ATTGATGATT	GGGATCCGGT	AGAAGATGAA	AATATGCTGG	ATAACATTGT	CAAAACTATA	13800
AATGATAACT	GTAATAAAGA	TAATAAAGGG	AATAAAATTA	ACAATTTCTG	GGGACTAGCA	13860
CTTAAGAAGT	ATCAAGTCCT	TAAAATCAGA	TCTATAACAA	GTGATTCTGA	TGATAATGAT	13920
AGACTAGATG	CTAATACAAG	TGGTTTGACA	CTTCTCAAG	GAGGGAATTA	TCTATCGCAT	13980
CAATTGAGAT	TATTCGGAAT	CAACAGCACT	AGTTGTCTGA	AAGCTCTTGA	GTTATCACAA	14040
ATTTTAATGA	AGGAAGTCAA	TAAAGACAAG	GACAGGCTCT	TCCTGGGAGA	AGGAGCAGGA	14100
GCTATGCTAG	CATGTTATGA	TGCCACATTA	GGACCTGCAG	TTAATTATTA	TAATTCAGGT	14160
TTGAATATAA	CAGATGTAAT	TGGTCAACGA	GAATTGAAAA	TATTTCTTTC	AGAGGTATCA	14220
TTAGTAGGTA	AAAAATTAGG	AAATGTGACA	CAGATTCTTA	ACAGGGTAAA	AGTACTGTTC	14280
AATGGGAATC	CTAATTCAAC	ATGGATAGGA	AATATGGAAT	GTGAGAGCTT	AATATGGAGT	14340
GAATTAAATG	ATAAGTCCAT	TGGATTAGTA	CATTGTGATA	TGGAAGGAGC	TATCGGTAAA	14400
TCAGAAGAAA	CTGTTCTACA	TGAACATTAT	AGTGTTATAA	GAATTACATA	CTTGATTGGG	14460

GATGATGATG TTGTTTTAGT TTCCAAAATT ATACCTACAA TCACTCCGAA TTGGTCTAGA 14520  
 ATACTTTATC TATATAAATT ATATTGGAAA GATGTAAGTA TAATATCACT CAAAACCTTCT 14580  
 AATCCTGCAT CAACAGAATT ATATCTAATT TCGAAAGATG CATATTGTAC TATAATGGAA 14640  
 CCTAGTGAAA TTGTTTTATC AAAACTTAAA AGATTGTCAC TCTTGGAAGA AAATAATCTA 14700  
 TTAAAATGGA TCATTTTATC AAAGAAGAGG AATAATGAAT GGTTACATCA TGAAATCAAA 14760  
 GAAGGAGAAA GAGATTATGG AATCATGAGA CCATATCATA TGGCACTACA AATCTTTGGA 14820  
 TTTCAAATCA ATTTAAATCA TCTGGCGAAA GAATTTTTTAT CAACCCCGA TCTGACTAAT 14880  
 ATCAACAATA TAATCCAAAG TTTTCAGCGA ACAATAAAGG ATGTTTTATT TGAATGGATT 14940  
 AATATAACTC ATGATGATAA GAGACATAAA TTAGGCGGAA GATATAACAT ATTCCCCTG 15000  
 AAAAATAAGG GAAAGTTAAG ACTGCTATCG AGAAGACTAG TATTAAGTTG GATTTTCATTA 15060  
 TCATTATCGA CTCGATTACT TACAGGTCGC TTTCCTGATG AAAAATTTGA ACATAGAGCA 15120  
 CAGACTGGAT ATGTATCATT AGCTGATACT GATTTAGAAT CATTAAAGTT ATTGTCGAAA 15180  
 AACATCATTAGA AATTACAG AGAGTGTATA GGATCAATAT CATATTGGTT TCTAACCAAA 15240  
 GAAGTTAAAA TACTTATGAA ATTGATCGGT GGTGCTAAAT TATTAGGAAT TCCCAGACAA 15300  
 TATAAAGAAC CCGAAGACCA GTTATTAGAA AACTACAATC AACATGATGA ATTTGATATC 15360  
 GATTAAAACA TAAATACAAT GAAGATATAT CCTAACCTTT ATCTTTAAGC CTAGGAATAG 15420  
 ACAAAAAGTA AGAAAAACAT GTAATATATA TATACCAAAC AGAGTTCTTC TCTTGTTTGG 15480  
 TGGGTCGGCA TGGCATCTCC ACCTCCTCGC GGTCCGGACC TGGGCATCCG AAGGAGGACG 15540  
 CACGTCCACT CGGATGGCTA AGGGAGAGCC TGCAGTAGCA TAACCCCTTG GGGCCTCTAA 15600  
 ACGGGTCTTG AGGGGTTTTT TGCTGAAAGG AGGAACTATA TACGCGTCGA CGGGCCCCGC 15660  
 GCTCAC 15666

## (2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGATTTGCGC GCAATTTAAA TCATCTGG

28

## (2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 70 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCCAGGTCGG ACCGCGAGGA GGTGGAGATG CCATGCCAGC CCACCAAAC AAGAGAAGAA 60  
CTCTGTTTGG 70

## (2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 45 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGCCCCTCGA CGCGTAATAC GACTCACTAT AGGACCAAAC AAGAG 45

## (2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 16 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CGGCATCACG TGCTAC 16

## (2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 6843 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATGTTGAGCC TATTTGATAC ATTTAATGCA CGTAGGCAAG AAAACATAAC AAAATCAGCC 60  
 GGTGGAGCTA TCATTCTTGG ACAGAAAAAT ACTGTCTCTA TATTCGCCCT TGGACCGACA 120  
 ATAAGTGATG ATAATGAGAA AATGACATTA GCTCTTCTAT TTCTATCTCA TTCACTAGAT 180  
 AATGAGAAAC AACATGCACA AAGGGCAGGG TTCTTGGTGT CTTTATTGTC AATGGCTTAT 240  
 GCCAATCCAG AGCTCTACCT AACAACAAAT GGAAGTAATG CAGATGTCAA GTATGTCATA 300  
 TACATGATTG AGAAAGATCT AAAACGGCAA AAGTATGGAG GATTTGTGGT TAAGACGAGA 360  
 GAGATGATAT ATGAAAAGAC AACTGATTGG ATATTTGGAA GTGACCTGGA TTATGATCAG 420  
 GAAACTATGT TGCAGAACGG CAGGAACAAT TCAACAATTG AAGACCTTGT CCACACATTT 480

GGGTATCCAT	CATGTTTAGG	AGCTCTTATA	ATACAGATCT	GGATAGTTCT	GGTCAAAGCT	540
ATCACTAGTA	TCTCAGGGTT	AAGAAAAAGC	TTTTTCACCC	GATTGGAAGC	TTTCAGACAA	600
GATGGAACAG	TGCAGGCAGG	GCTGGTATTG	AGCGGTGACA	CAGTGGATCA	GATTGGGTCA	660
ATCATGCGGT	CTCAACAGAG	CTTGGTAACT	CTTATGGTTG	AAACATTAAT	AACAATGAAT	720
ACCAGCAGAA	ATGACCTCAC	AACCATAGAA	AAGAATATAC	AAATTGTTGG	CAACTACATA	780
AGAGATGCAG	GTCTCGCTTC	ATTCTTCAAT	ACAATCAGAT	ATGGAATTGA	GACCAGAATG	840
GCAGCTTTGA	CTCTATCCAC	TCTCAGACCA	GATATCAATA	GATTAAAAGC	TTTGATGGAA	900
CTGTATTTAT	CAAAGGGACC	ACGCGCTCCT	TTCATCTGTA	TCCTCAGAGA	TCCTATACAT	960
GGTGAGTTTCG	CACCAGGCAA	CTATCCTGCC	ATATGGAGCT	ATGCAATGGG	GGTGGCAGTT	1020
GTACAAAATA	GAGCCATGCA	ACAGTATGTG	ACGGGAAGAT	CATATCTAGA	CATTGATATG	1080
TTCCAGCTAG	GACAAGCAGT	AGCACGTGAT	GCCGAAGCTC	AAATGAGCTC	AACACTGGAA	1140
GATGAAC TTG	GAGTGACACA	CGAATCTAAA	GAAAGCTTGA	AGAGACATAT	AAGGAACATA	1200
AACAGTTCAG	AGACATCTTT	CCACAAACCG	ACAGGTGGAT	CAGCCATAGA	GATGGCAATA	1260
GATGAAGAGC	CAGAACAATT	CGAACATAGA	GCAGATCAAG	AACAAAATGG	AGAACCTCAA	1320
TCATCCATAA	TTCAATATGC	CTGGGCAGAA	GGAAATAGAA	GCGATGATCA	GACTGAGCAA	1380
GCTACAGAAT	CTGACAATAT	CAAGACCGAA	CAACAAAACA	TCAGAGACAG	ACTAAACAAG	1440
AGACTCAACG	ACAAGAAGAA	ACAAAGCAGT	CAACCACCCA	CTAATCCCAC	AAACAGAACA	1500
AACCAGGACG	AAATAGATGA	TCTGTTTAAAC	GCATTTGGAA	GCAACTAAGT	CGACGATCCG	1560
GCTGCTAACA	AAGCCCCGAA	GGAAGCTGAG	TTGGCTGCTG	CCACCGCTGA	GCAATAACTA	1620
GCATAACCCC	TTGGGGCCTC	TAAACGGGTC	TTGAGGGGTT	TTTTGCTGAA	AGGAGGAACT	1680
ATATCCGGAT	CGAGATCAAT	TCTGTGAGCG	TATGGCAAAC	GAAGGAAAAA	TAGTTATAGT	1740
AGCCGCACTC	GATGGGACAT	TTCAACGTAA	ACCGTTTAAT	AATATTTTGA	ATCTTATTCC	1800
ATTATCTGAA	ATGGTG GTAA	AACTAACTGC	TGTGTGTATG	AAATGCTTTA	AGGAGGCTTC	1860
CTTTTCTAAA	CGATTGGGTG	AGGAAACCGA	GATAGAAATA	ATAGGAGGTA	ATGATATGTA	1920
TCAATCGGTG	TGTAGAAAGT	GTTACATCGA	CTCATAATAT	TATATTTTTT	ATCTAAAAAA	1980
CTAAAAATAA	ACATTGATTA	AATTTTAATA	TAATACTTAA	AAATGGATGT	TGTGTCGTTA	2040
GATAAACCGT	TTATGTATTT	TGAGGAAATT	GATAATGAGT	TAGATTACGA	ACCAGAAAGT	2100
GCAATGAGG	TCGCAAAAAA	ACTGCCGTAT	CAAGGACAGT	TAAAAC TATT	ACTAGGAGAA	2160
TTATTTTTTTC	TTAGTAAGTT	ACAGCGACAC	GGTATATTAG	ATGGTGCCAC	CGTAGTGTAT	2220
ATAGGATCTG	CTCCCGGTAC	ACATATACGT	TATTTGAGAG	ATCATTTCTA	TAATTTAGGA	2280
GTGATCATCA	AATGGATGCT	AATTGACGGC	CGCCATCATG	ATCCTATTTT	AAATGGATTG	2340
CGTGATGTGA	CTCTAGTGAC	TCGGTTCGTT	GATGAGGAAT	ATCTACGATC	CATCAAAAAA	2400
CAACTGCATC	CTTCTAAGAT	TATTTTAATT	TCTGATGTGA	GATCCAAACG	AGGAGGAAAT	2460
GAACCTAGTA	CGGCGGATTT	ACTAAGTAAT	TACGCTCTAC	AAAATGTCAT	GATTAGTATT	2520
TTAAACCCCG	TGGCGTCTAG	TCTTAAATGG	AGATGCCCGT	TTCCAGATCA	ATGGATCAAG	2580

GACTTTTATA	TCCCACACGG	TAATAAAATG	TTACAACCTT	TTGCTCCTTC	ATATTCAGGG	2640
CCGTCGTTTT	ACAACGTCGT	GACTGGGAAA	ACCCTGGCGT	TACCCAACCT	AATCGCCTTG	2700
CAGCACATCC	CCCTTTCGCC	AGCTGGCGTA	ATAGCGAAGA	GGCCCGCACC	GATCGCCCTT	2760
CCCAACAGTT	GCGCAGCCTG	AATGGCGAAT	GGCGCGACGC	GCCCTGTAGC	GGCGCATTA	2820
GCGCGGCGGG	TGTGGTGGTT	ACGCGCAGCG	TGACCGCTAC	ACTTGCCAGC	GCCCTAGCGC	2880
CCGCTCCTTT	CGCTTTCCTC	CCTTCCTTTC	TCGCCACGTT	CGCCGGCTTT	CCCCGTCAAG	2940
CTCTAAATCG	GGGGCTCCCT	TTAGGGTTCC	GATTTAGTGC	TTTACGGCAC	CTCGACCCCA	3000
AAAACTTGA	TTAGGGTGAT	GGTTCACGTA	GTGGGCCATC	GCCCTGATAG	ACGGTTTTTC	3060
GCCCTTTGAC	GTTGGAGTCC	ACGTTCTTTA	ATAGTGGACT	CTTGTTCCAA	ACTGGAACAA	3120
CACTCAACCC	TATCTCGGTC	TATTCCTTTT	ATTTATAAGG	GATTTTGCCG	ATTTTCGGCT	3180
ATTGGTTAAA	AAATGAGCTG	ATTTAACAAA	AATTTAACGC	GAATTTTAAC	AAAATATTAA	3240
CGTTTACAAT	TTCCCAGGTG	GCACTTTTCG	GGGAAATGTG	CGCGGAACCC	CTATTTGTTT	3300
ATTTTTCTAA	ATACATTCAA	ATATGTATCC	GCTCATGAGA	CAATAACCCCT	GATAAATGCT	3360
TCAATAATAT	TGAAAAAGGA	AGAGTATGAG	TATTCAACAT	TTCCGTGTCT	CCCTTATTCC	3420
CTTTTTTGCG	GCATTTTGCC	TTCTGTTTTT	TGCTCACCCA	GAAACGCTGG	TGAAAGTAAA	3480
AGATGCTGAA	GATCAGTTGG	GTGCACGAGT	GGGTACATC	GAACGGATC	TCAACAGCGG	3540
TAAGATCCTT	GAGAGTTTTT	GCCCCGAAGA	ACGTTTTCCA	ATGATGAGCA	CTTTTAAAGT	3600
TCTGCTATGT	GGCGCGGTAT	TATCCCGTAT	TGACGCCGGG	CAAGAGCAAC	TCGGTCGCCC	3660
CATACACTAT	TCTCAGAAAT	ACTTGGTGTA	GTACTACCA	GTACAGAAA	AGCATCTTAC	3720
GGATGGCATG	ACAGTAAGAG	AATTATGCAG	TGCTGCCATA	ACCATGAGTG	ATAACACTGC	3780
GGCCAACTTA	CTTCTGACAA	CGATCGGAGG	ACCGAAGGAG	CTAACCGCTT	TTTTGCACAA	3840
CATGGGGGAT	CATGTAATC	GCCTTGATCG	TTGGGAACCG	GAGCTGAATG	AAGCCATACC	3900
AAACGACGAG	CGTGACACCA	CGATGCCTGT	AGCAATGGCA	ACAACGTTGC	GCAAACATTT	3960
AACTGGCGAA	CTACTTACTC	TAGCTTCCCG	GCAACAATTA	ATAGACTGGA	TGGAGGCGGA	4020
TAAAGTTGCA	GGACCACTTC	TGCGCTCGGC	CCTTCCGGCT	GGCTGGTTTA	TTGCTGATAA	4080
ATCTGGAGCC	GGTGAGCGTG	GGTCTCGCGG	TATCATTGCA	GCACTGGGGC	CAGATGGTAA	4140
GCCCTCCCGT	ATCGTAGTTA	TCTACACGAC	GGGGAGTCAG	GCAACTATGG	ATGAACGAAA	4200
TAGACAGATC	GCTGAGATAG	GTGCCTCACT	GATTAAGCAT	TGGTAACTGT	CAGACCAAGT	4260
TTACTCATAT	ATACTTTAGA	TTGATTTAAA	ACTTCATTTT	TAATTTAAAA	GGATCTAGGT	4320
GAAGATCCTT	TTTGATAATC	TCATGACCAA	AATCCCTTAA	CGTGAGTTTT	CGTTCCACTG	4380
AGCGTCAGAC	CCCGTAGAAA	AGATCAAAGG	ATCTTCTTGA	GATCCTTTTT	TTCTGCGCGT	4440
AATCTGCTGC	TTGCAAACAA	AAAAACCACC	GCTACCAGCG	GTGGTTTGTT	TGCCGGATCA	4500
AGAGCTACCA	ACTCTTTTTT	CGAAGGTAAC	TGGCTTCAGC	AGAGCGCAGA	TACCAAATAC	4560
TGTCCTTCTA	GTGTAGCCGT	AGTTAGGCCA	CCACTTCAAG	AACTCTGTAG	CACCGCCTAC	4620
ATACCTCGCT	CTGCTAATCC	TGTTACCAGT	GGCTGCTGCC	AGTGGCGATA	AGTCGTGTCT	4680

TACCGGGTTG	GACTCAAGAC	GATAGTTACC	GGATAAGGCG	CAGCGGTCGG	GCTGAACGGG	4740
GGGTTCGTGC	ACACAGCCCA	GCTTGGAGCG	AACGACCTAC	ACCGAACTGA	GATACCTACA	4800
GCGTGAGCTA	TGAGAAAGCG	CCACGCTTCC	CGAAGGGAGA	AAGGCGGACA	GGTATCCGGT	4860
AAGCGGCAGG	GTCGGAACAG	GAGAGCGCAC	GAGGGAGCTT	CCAGGGGGAA	ACGCCTGGTA	4920
TCTTTATAGT	CCTGTCGGGT	TTGCCACCT	CTGACTTGAG	CGTCGATTTT	TGTGATGCTC	4980
GTCAGGGGGG	CGGAGCCTAT	GGAAAAACGC	CAGCAACGCG	GCCTTTTTTAC	GGTTCCTGGC	5040
CTTTTGCTGG	CCTTTTGCTC	ACATGTTCTT	TCCTGCGTTA	TCCCCTGATT	CTGTGGATAA	5100
CCGTATTACC	GCCTTTGAGT	GAGCTGATAC	CGCTCGCCGC	AGCCGAACGA	CCGAGCGCAG	5160
CGAGTCAGTG	AGCGAGGAAG	CGGAAGAGCG	CCCAATACGC	AAACCGCCTC	TCCCCGCGCG	5220
TTGGCCGATT	CATTAATGCA	GCTGGCACGA	CAGGTTTCCC	GACTGGAAAG	CGGGCAGTGA	5280
GCGCAACGCA	ATTAATGTGA	GTTAGCTCAC	TCATTAGGCA	CCCCAGGCTT	TACACTTTAT	5340
GCTTCCGGCT	CGTATGTTGT	GTGGAATTGT	GAGCGGATAA	CAATTTTACA	CAGGAAACAG	5400
CTATGACCAT	GATTACGCCA	AGCTTTTGCG	ATCAATAAAT	GGATCACAAC	CAGTATCTCT	5460
TAACGATGTT	CTTCGCAGAT	GATGATTTCAT	TTTTTAAGTA	TTTGGCTAGT	CAAGATGATG	5520
AATCTTCATT	ATCTGATATA	TTGCAAATCA	CTCAATATCT	AGACTTTCTG	TTATTATTAT	5580
TGATCCAATC	AAAAAATAAA	TTAGAAGCCG	TGGGTCATTG	TTATGAATCT	CTTTCAGAGG	5640
AATACAGACA	ATTGACAAAA	TTACAGACT	TTCAAGATTT	TAAAAAACTG	TTTAAACAAGG	5700
TCCCTATTGT	TACAGATGGA	AGGGTCAAAC	TTAATAAAGG	ATATTTGTTC	GACTTTGTGA	5760
TTAGTTTGAT	GCGATTCAAA	AAAGAATCCT	CTCTAGCTAC	CACCGCAATA	GATCCTGTTA	5820
GATACATAGA	TCCTCGTCGC	AATATCGCAT	TTTCTAACGT	GATGGATATA	TTAAAGTCGA	5880
ATAAAGTGAA	CAATAATTAA	TTCTTTATTG	TCATCATGAA	CGGCGGACAT	ATTCAGTTGA	5940
TAATCGGCCC	CATGTTTTCA	GGTAAAAGTA	CAGAATTAAT	TAGACGAGTT	AGACGTTATC	6000
AAATAGCTCA	ATATAAATGC	GTGACTATAA	AATATTCTAA	CGATAATAGA	TACGGAACGG	6060
GACTATGGAC	GCATGATAAG	AATAATTTTG	AAGCATTGGA	AGCAACTAAA	CTATGTGATG	6120
TCTTGGAATC	AATTACAGAT	TTCTCCGTGA	TAGGTATCGA	TGAAGGACAG	TTCTTTCCAG	6180
ACATTGTTGA	ATTGATCTCG	ATCCCGCGAA	ATTAATACGA	CTCACTATAG	GGAGACCACA	6240
ACGGTTTCCC	TCTAGCGGGA	TCAATTCCGC	CCCTCTCCCT	CCCCCCCCC	TAACGTTACT	6300
GGCCGAAGCC	GCTTGGAATA	AGGCCGGTGT	GCGTTTGTCT	ATATGTTATT	TTCCACCATA	6360
TTGCCGTCTT	TTGGCAATGT	GAGGGCCCGG	AAACCTGGCC	CTGTCTTCTT	GACGAGCATT	6420
CCTAGGGGTC	TTTCCCCTCT	CGCCAAAGGA	ATGCAAGGTC	TGTTGAATGT	CGTGAAGGAA	6480
GCAGTTCCTC	TGGAAGCTTC	TTGAAGACAA	ACAACGTCTG	TAGCGACCCT	TTGCAGGCAG	6540
CGGAACCCCC	CACCTGGCGA	CAGGTGCCTC	TGCGGCCAAA	AGCCACGTGT	ATAAGATACA	6600
CCTGCAAAGG	CGGCACAACC	CCAGTGCCAC	GTTGTGAGTT	GGATAGTTGT	GGAAAGAGTC	6660
AAATGGCTCT	CCTCAAGCGT	ATTCAACAAG	GGGCTGAAGG	ATGCCCAGAA	GGTACCCCAT	6720
TGTATGGGAT	CTGATCTGGG	GCCTCGGTGC	ACATGCTTTA	CATGTGTTTA	GTCGAGGTTA	6780

(2) INFORMATION FOR SEQ ID NO:21:

(A) LENGTH: 7107 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ATGGAAAGCG	ACGCTAAAAA	CTATCAAATC	ATGGATTCTT	GGGAAGAGGA	ATCAAGAGAT	60
AAATCAACTA	ATATCTCCTC	GGCCCTCAAC	ATCATTGAAT	TCATACTCAG	CACCGACCCC	120
CAAGAAGACT	TATCGGAAAA	CGACACAATC	AACACAAGAA	CCCAGCAACT	CAGTGCCACC	180
ATCTGTCAAC	CAGAAATCAA	ACCAACAGAA	ACAAGTGAGA	AAGATAGTGG	ATCAACTGAC	240
AAAAATAGAC	AGTCCGGGTC	ATCACACGAA	TGTACAACAG	AAGCAAAAGA	TAGAAATATT	300
GATCAGGAAA	CTGTACAGAG	AGGACCTGGG	AGAAGAAGCA	GCTCAGATAG	TAGAGCTGAG	360
ACTGTGGTCT	CTGGAGGAAT	CCCCAGAAGC	ATCACAGATT	CTAAAAATGG	AACCCAAAAAC	420
ACGGAGGATA	TTGATCTCAA	TGAAATTAGA	AAGATGGATA	AGGACTCTAT	TGAGGGGAAA	480
ATGCGACAAT	CTGCAAATGT	TCCAAGCGAG	ATATCAGGAA	GTGATGACAT	ATTTACAACA	540
GAACAAAGTA	GAAACAGTGA	TCATGGAAGA	AGCCTGGAAT	CTATCAGTAC	ACCTGATACA	600
AGATCAATAA	GTGTTGTTAC	TGCTGCAACA	CCAGATGATG	AAGAAGAAAT	ACTAATGAAA	660
AATAGTAGGA	CAAAGAAAAG	TTCTTCAACA	CATCAAGAAG	ATGACAAAAG	AATTAAAAAA	720
GGGGGAAAAG	GGAAAGACTG	GTTTAAGAAA	TCAAAGATA	CCGACAACCA	GATACCAACA	780
TCAGACTACA	GATCCACATC	AAAAGGGCAG	AAGAAAATCT	CAAAGACAAC	AACCACCAAC	840
ACCGACACAA	AGGGGCAAAC	AGAAATACAG	ACAGAATCAT	CAGAAACACA	ATCCTCATCA	900
TGGAATCTCA	TCATCGACAA	CAACACCGAC	CGGAACGAAC	AGACAAGCAC	AACTCCTCCA	960
ACAACAAC TT	CCAGATCAAC	TTATACAAAA	GAATCGATCC	GAACAAACTC	TGAATCCAAA	1020
CCCAAGACAC	AAAAGACAAA	TGGAAAGGAA	AGGAAGGATA	CAGAAGAGAG	CAATCGATT T	1080
ACAGAGAGGG	CAATTACTCT	ATTGCAGAAT	CTTGGTGTA A	TTCAATCCAC	ATCAAAACTA	1140
GATTTATATC	AAGACAAACG	AGTTGTATGT	GTAGCAAATG	TACTAAACAA	TGTAGATACT	1200
GCATCAAAGA	TAGATTTCCT	GGCAGGATTA	GTCATAGGGG	TTTCAATGGA	CAACGACACA	1260
AAATTAACAC	AGATACAAAA	TGAAATGCTA	AACCTCAAAG	CAGATCTAAA	GAAAATGGAC	1320
GAATCACATA	GAAGATTGAT	AGAAAATCAA	AGAGAACAAC	TGTCATTGAT	CACGTCACTA	1380
ATTTCAAATC	TCAAAAATTAT	GACTGAGAGA	GGAGGAAAGA	AAGACCAAAA	TGAATCCAAT	1440
GAGAGAGTAT	CCATGATCAA	AACAAAATTG	AAAGAAGAAA	AGATCAAGAA	GACCAGGTTT	1500
GACCCACTTA	TGGAGGCACA	AGGCATTGAC	AAGAATATAC	CCGATCTATA	TCGACATGCA	1560



GGAGATACAC	TAGAGAACGA	TGTACAAGTT	AAATCAGAGA	TATTAAGTTC	ATACAATGAG	1620
TCAAATGCAA	CAAGACTAAT	ACCCAAAAAA	GTGAGCAGTA	CAATGAGATC	ACTAGTTGCA	1680
GTCATCAACA	ACAGCAATCT	CTCACAAAGC	ACAAAACAAT	CATACATAAA	CGAACTCAAA	1740
CGTTGCAAAA	ATGATGAAGA	AGTATCTGAA	TTAATGGACA	TGTTCAATGA	AGATGTCAAC	1800
AATTGCCAAT	GAGTCGACGA	TCCGGCTGCT	AACAAAGCCC	GAAAGGAAGC	TGAGTTGGCT	1860
GCTGCCACCG	CTGAGCAATA	ACTAGCATAA	CCCCTTGGGG	CCTCTAAACG	GGTCTTGAGG	1920
GGTTTTTTGC	TGAAAGGAGG	AACTATATCC	GGATCGAGAT	CAATTCTGTG	AGCGTATGGC	1980
AAACGAAGGA	AAAATAGTTA	TAGTAGCCGC	ACTCGATGGG	ACATTTCAAC	GTAAACCGTT	2040
TAATAATATT	TTGAATCTTA	TTCCATTATC	TGAAATGGTG	GTAAAACTAA	CTGCTGTGTG	2100
TATGAAATGC	TTTAAGGAGG	CTTCCTTTTC	TAAACGATTG	GGTGAGGAAA	CCGAGATAGA	2160
AATAATAGGA	GGTAATGATA	TGTATCAATC	GGTGTGTAGA	AAGTGTTACA	TCGACTCATA	2220
ATATTATATT	TTTTATCTAA	AAAACATAAA	ATAAACATTG	ATTAAATTTT	AATATAATAC	2280
TTAAAAATGG	ATGTTGTGTC	GTTAGATAAA	CCGTTTATGT	ATTTTGAGGA	AATTGATAAT	2340
GAGTTAGATT	ACGAACCAGA	AAGTGCAAAT	GAGGTCGCAA	AAAAACTGCC	GTATCAAGGA	2400
CAGTTAAAC	TATTACTAGG	AGAATTATTT	TTTCTTAGTA	AGTTACAGCG	ACACGGTATA	2460
TTAGATGGTG	CCACCGTAGT	GTATATAGGA	TCTGCTCCCG	GTACACATAT	ACGTTATTTG	2520
AGAGATCATT	TCTATAATTT	AGGAGTGATC	ATCAAATGGA	TGCTAATTGA	CGGCCGCCAT	2580
CATGATCCTA	TTTTTAAATGG	ATTGCGTGAT	GTGACTCTAG	TGACTCGGTT	CGTTGATGAG	2640
GAATATCTAC	GATCCATCAA	AAAACAAC TG	CATCCTTCTA	AGATTATTTT	AATTTCTGAT	2700
GTGAGATCCA	AACGAGGAGG	AAATGAACCT	AGTACGGCGG	ATTTACTAAG	TAATTACGCT	2760
CTACAAAATG	TCATGATTAG	TATTTTAAAC	CCCGTGGCGT	CTAGTCTTAA	ATGGAGATGC	2820
CCGTTTCCAG	ATCAATGGAT	CAAGGACTTT	TATATCCAC	ACGGTAATAA	AATGTTACAA	2880
CCTTTTGCTC	CTTCATATTC	AGGGCCGTCG	TTTTACAACG	TCGTGACTGG	GAAAACCCTG	2940
GCGTTACCCA	ACTTAATCGC	CTTGCAGCAC	ATCCCCCTTT	CGCCAGCTGG	CGTAATAGCG	3000
AAGAGGCCCG	CACCGATCGC	CCTTCCCAAC	AGTTGCGCAG	CCTGAATGGC	GAATGGCGCG	3060
ACGCGCCCTG	TAGCGGCGCA	TTAAGCGCGG	CGGGTGTGGT	GGTTACGCGC	AGCGTGACCG	3120
CTACACTTGC	CAGCGCCCTA	GCGCCCGCTC	CTTTCGCTTT	CTTCCCTTCC	TTTCTCGCCA	3180
CGTTCGCCGG	CTTTCCCCGT	CAAGCTCTAA	ATCGGGGGCT	CCCTTTAGGG	TTCCGATTTA	3240
GTGCTTTACG	GCACCTCGAC	CCCAAAAAAC	TTGATTAGGG	TGATGGTTCA	CGTAGTGGGC	3300
CATCGCCCTG	ATAGACGGTT	TTTCGCCCTT	TGACGTTGGA	GTCCACGTTC	TTTAATAGTG	3360
GACTCTTGTT	CCAAACTGGA	ACAACACTCA	ACCCTATCTC	GGTCTATTCT	TTTGATTTAT	3420
AAGGGATTTT	GCCGATTTTCG	GCCTATTGGT	TAAAAAATGA	GCTGATTTAA	CAAAAATTTA	3480
ACGCGAATTT	TAACAAAATA	TTAACGTTTA	CAATTTCCCA	GGTGGCACTT	TTCGGGGAAA	3540
TGTGCGCGGA	ACCCCTATTT	GTTTATTTTT	CTAAATACAT	TCAAATATGT	ATCCGCTCAT	3600
GAGACAATAA	CCCTGATAAA	TGCTTCAATA	ATATTGAAAA	AGGAAGAGTA	TGAGTATTCA	3660

ACATTTCCGT	GTCGCCCTTA	TTCCCTTTTT	TGCGGCATTT	TGCCTTCCTG	TTTTTGCTCA	3720
CCCAGAAACG	CTGGTGAAAG	TAAAAGATGC	TGAAGATCAG	TTGGGTGCAC	GAGTGGGTTA	3780
CATCGAACTG	GATCTCAACA	GCGGTAAGAT	CCTTGAGAGT	TTTCGCCCCG	AAGAACGTTT	3840
TCCAATGATG	AGCACTTTTA	AAGTTCTGCT	ATGTGGCGCG	GTATTATCCC	GTATTGACGC	3900
CGGGCAAGAG	CAACTCGGTC	GCCGCATACA	CTATTCTCAG	AATGACTTGG	TTGAGTACTC	3960
ACCAGTCACA	GAAAAGCATC	TTACGGATGG	CATGACAGTA	AGAGAATTAT	GCAGTGCTGC	4020
CATAACCATG	AGTGATAACA	CTGCGGCCAA	CTTACTTCTG	ACAACGATCG	GAGGACCGAA	4080
GGAGCTAACC	GCTTTTTTGC	ACAACATGGG	GGATCATGTA	ACTCGCCTTG	ATCGTTGGGA	4140
ACCGGAGCTG	AATGAAGCCA	TACCAAACGA	CGAGCGTGAC	ACCACGATGC	CTGTAGCAAT	4200
GGCAACAACG	TTGCGCAAAC	TATTAACCTG	CGAACTACTT	ACTCTAGCTT	CCCGGCAACA	4260
ATTAATAGAC	TGGATGGAGG	CGGATAAAGT	TGCAGGACCA	CTTCTGCGCT	CGGCCCTTCC	4320
GGCTGGCTGG	TTTATTGCTG	ATAAATCTGG	AGCCGGTGAG	CGTGGGTCTC	GCGGTATCAT	4380
TGCAGCACTG	GGGCCAGATG	GTAAGCCCTC	CCGTATCGTA	GTTATCTACA	CGACGGGGAG	4440
TCAGGCAACT	ATGGATGAAC	GAAATAGACA	GATCGCTGAG	ATAGGTGCCT	CACTGATTAA	4500
GCATTGGTAA	CTGTCAGACC	AAGTTTACTC	ATATATACTT	TAGATTGATT	TAAAACTTCA	4560
TTTTTAATTT	AAAAGGATCT	AGGTGAAGAT	CCTTTTTGAT	AATCTCATGA	CCAAAATCCC	4620
TTAACGTGAG	TTTTTCGTTCC	ACTGAGCGTC	AGACCCCGTA	GAAAAGATCA	AAGGATCTTC	4680
TTGAGATCCT	TTTTTTCTGC	GCGTAATCTG	CTGCTTGCAA	ACAAAAAAC	CACCGCTACC	4740
AGCGGTGGTT	TGTTTGCCGG	ATCAAGAGCT	ACCAACTCTT	TTTCCGAAGG	TAAGTGCTT	4800
CAGCAGAGCG	CAGATACCAA	ATACTGTCCT	TCTAGTGTAG	CCGTAGTTAG	GCCACCACTT	4860
CAAGAACTCT	GTAGACCCGC	CTACATACCT	CGCTCTGCTA	ATCCTGTTAC	CAGTGGCTGC	4920
TGCCAGTGGC	GATAAGTCGT	GTCTTACCGG	GTTGGACTCA	AGACGATAGT	TACCGGATAA	4980
GGCGCAGCGG	TCGGGCTGAA	CGGGGGGTTT	GTGCACACAG	CCCAGCTTGG	AGCGAACGAC	5040
CTACACCGAA	CTGAGATACC	TACAGCGTGA	GCTATGAGAA	AGCGCCACGC	TTCCCGAAGG	5100
GAGAAAGGCG	GACAGGTATC	CGGTAAGCGG	CAGGGTCGGA	ACAGGAGAGC	GCACGAGGGA	5160
GCTTCCAGGG	GGAAACGCCT	GGTATCTTTA	TAGTCCTGTC	GGGTTTCGCC	ACCTCTGACT	5220
TGAGCGTCGA	TTTTTGTGAT	GCTCGTCAGG	GGGGCGGAGC	CTATGGAAAA	ACGCCAGCAA	5280
CGCGGCCTTT	TTACGGTTCC	TGGCCTTTTG	CTGGCCTTTT	GCTCACATGT	TCTTTCCTGC	5340
GTTATCCCCT	GATTCTGTGG	ATAACCGTAT	TACCGCCTTT	GAGTGAGCTG	ATACCGCTCG	5400
CCGCAGCCGA	ACGACCGAGC	GCAGCGAGTC	AGTGAGCGAG	GAAGCGGAAG	AGCGCCCAAT	5460
ACGCAAACCG	CCTCTCCCCG	CGCGTTGGCC	GATTCATTAA	TGCAGCTGGC	ACGACAGGTT	5520
TCCCGACTGG	AAAGCGGGCA	GTGAGCGCAA	CGCAATTAAT	GTGAGTTAGC	TCACTCATTA	5580
GGCACCCAG	GCTTTTAACT	TTATGCTTCC	GGCTCGTATG	TTGTGTGGAA	TTGTGAGCGG	5640
ATAACAATTT	CACACAGGAA	ACAGCTATGA	CCATGATTAC	GCCAAGCTTT	TGCGATCAAT	5700
AAATGGATCA	CAACCAAGTAT	CTCTTAACGA	TGTTCTTCGC	AGATGATGAT	TCATTTTTTA	5760

AGTATTTGGC TAGTCAAGAT GATGAATCTT CATTATCTGA TATATTGCAA ATCACTCAAT 5820  
 ATCTAGACTT TCTGTTATTA TTATTGATCC AATCAAAAAA TAAATTAGAA GCCGTGGGTC 5880  
 ATTGTTATGA ATCTCTTTCA GAGGAATACA GACAATTGAC AAAATTCACA GACTTTCAAG 5940  
 ATTTTAAAAA ACTGTTTAAC AAGGTCCCTA TTGTTACAGA TGGAAGGGTC AAACCTAATA 6000  
 AAGGATATTT GTTCGACTTT GTGATTAGTT TGATGCGATT CAAAAAGAA TCCTCTCTAG 6060  
 CTACCACCGC AATAGATCCT GTTAGATACA TAGATCCTCG TCGCAATATC GCATTTTCTA 6120  
 ACGTGATGGA TATATTAAAG TCGAATAAAG TGAACAATAA TTAATTCTTT ATTGTCATCA 6180  
 TGAACGGCGG ACATATTCAG TTGATAATCG GCCCCATGTT TTCAGGTAAA AGTACAGAAT 6240  
 TAATTAGACG AGTTAGACGT TATCAAATAG CTCAATATAA ATGCGTGAAT ATAAAAATATT 6300  
 CTAACGATAA TAGATACGGA ACGGGACTAT GGACGCATGA TAAGAATAAT TTTGAAGCAT 6360  
 TGGAAGCAAC TAAACTATGT GATGTCTTGG AATCAATTAC AGATTTCTCC GTGATAGGTA 6420  
 TCGATGAAGG ACAGTTCTTT CCAGACATTG TTGAATTGAT CTCGATCCCG CGAAATTAAT 6480  
 ACGACTCACT ATAGGGAGAC CACAACGGTT TCCCTCTAGC GGGATCAATT CCGCCCCCTCT 6540  
 CCCTCCCCC CCCCTAACGT TACTGGCCGA AGCCGCTTGG AATAAGGCCG GTGTGCGTTT 6600  
 GTCTATATGT TATTTTCCAC CATATTGCCG TCTTTTGGCA ATGTGAGGGC CCGGAAACCT 6660  
 GGCCCTGTCT TCTTGACGAG CATTCCTAGG GGTCTTTCCC CTCTCGCAA AGGAATGCAA 6720  
 GGTCTGTTGA ATGTCGTGAA GGAAGCAGTT CCTCTGGAAG CTTCTTGAAG ACAAACAACG 6780  
 TCTGTAGCGA CCCTTTGCAG GCAGCGGAAC CCCCACCTG GCGACAGGTG CCTCTGCGGC 6840  
 CAAAAGCCAC GTGTATAAGA TACACCTGCA AAGGCGGCAC AACCCAGTG CCACGTTGTG 6900  
 AGTTGGATAG TTGTGGAAAG AGTCAAATGG CTCTCCTCAA GCGTATTCAA CAAGGGGCTG 6960  
 AAGGATGCCC AGAAGGTACC CCATTGTATG GGATCTGATC TGGGGCCTCG GTGCACATGC 7020  
 TTTACATGTG TTTAGTCGAG GTTAAAAAAC GTCTAGGCCC CCCGAACCAC GGGGACGTGG 7080  
 TTTTCCTTTG AAAAACACGA TAATACC 7107

## (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 12011 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ATGGACACTG AATCTAACAA TGGCACTGTA TCTGACATAC TCTATCCTGA GTGTCACCTT 60  
 AACTCTCCTA TCGTTAAAGG TAAAATAGCA CAATTACACA CTATTATGAG TCTACCTCAG 120  
 CCTTATGATA TGGATGACGA CTCAATACTA GTTATCACTA GACAGAAAAT AAAACTTAAT 180  
 AAATTGGATA AAAGACAACG ATCTATTAGA AGATTAAAAT TAATATTAAC TGAAAAAGTG 240  
 AATGACTTAG GAAAATACAC ATTTATCAGA TATCCAGAAA TGTCAAAAGA AATGTTCAAA 300

TTATATATAC	CTGGTATTAA	CAGTAAAGTG	ACTGAATTAT	TACTTAAAGC	AGATAGAACA	360
TATAGTCAAA	TGACTGATGG	ATTAAGAGAT	CTATGGATTA	ATGTGCTATC	AAAATTAGCC	420
TCAAAAAATG	ATGGAAGCAA	TTATGATCTT	AATGAAGAAA	TTAATAATAT	ATCGAAAGTT	480
CACACAACCT	ATAAATCAGA	TAAATGGTAT	AATCCATTCA	AAACATGGTT	TACTATCAAG	540
TATGATATGA	GAAGATTACA	AAAAGCTCGA	AATGAGATCA	CTTTTAATGT	TGGGAAGGAT	600
TATAACTTGT	TAGAAGACCA	GAAGAATTTT	TTATTGATAC	ATCCAGAATT	GGTTTTGATA	660
TTAGATAAAC	AAAAC TATAA	TGGTTATCTA	ATTACTCCTG	AATTAGTATT	GATGTATTGT	720
GACGTAGTCG	AAGGCCGATG	GAATATAAGT	GCATGTGCTA	AGTTAGATCC	AAAATTACAA	780
TCTATGTATC	AGAAAGGTAA	TAACCTGTGG	GAAGTGATAG	ATAAATTGTT	TCCAATTATG	840
GGAGAAAAGA	CATTTGATGT	GATATCGTTA	TTAGAACCAC	TTGCATTATC	CTTAATTCAA	900
ACTCATGATC	CTGTTAAACA	ACTAAGAGGA	GCTTTTTTTAA	ATCATGTGTT	ATCCGAGATG	960
GAATTAATAT	TTGAATCTAG	AGAATCGATT	AAGGAATTTT	TGAGTGTAGA	TTACATTGAT	1020
AAAATTTTAG	ATATATTTAA	TAAGTCTACA	ATAGATGAAA	TAGCAGAGAT	TTTCTCTTTT	1080
TTTAGAACAT	TTGGGCATCC	TCCATTAGAA	GCTAGTATTG	CAGCAGAAAA	GGTTAGAAAA	1140
TATATGTATA	TTGGAAAAACA	ATTAAAATTT	GACACTATTA	ATAAATGTCA	TGCTATCTTC	1200
TGTACAATAA	TAATTAACGG	ATATAGAGAG	AGGCATGGTG	GACAGTGGCC	TCCTGTGACA	1260
TTACCTGATC	ATGCACACGA	ATTCATCATA	AATGCTTACG	GTTCAAAC TC	TGCGATATCA	1320
TATGAAAATG	CTGTTGATTA	TTACCAGAGC	TTTATAGGAA	TAAAATTCAA	TAAATTCATA	1380
GAGCCTCAGT	TAGATGAGGA	TTTGACAATT	TATATGAAAG	ATAAAGCATT	ATCTCCAAAA	1440
AAATCAAATT	GGGACACAGT	TTATCCTGCA	TCTAATTTAC	TGTACCGTAC	TAACGCATCC	1500
AACGAATCAC	GAAGATTAGT	TGAAGTATTT	ATAGCAGATA	GTAAATTTGA	TCCTCATCAG	1560
ATATTGGATT	ATGTAGAATC	TGGGGACTGG	TTAGATGATC	CAGAATTTAA	TATTTCTTAT	1620
AGTCTTAAAG	AAAAAGAGAT	CAAACAGGAA	GGTAGACTCT	TTGCAAAAAT	GACATACAAA	1680
ATGAGAGCTA	CACAAGTTTT	ATCAGAGACA	CTACTTGCAA	ATAACATAGG	AAAATTCTTT	1740
CAAGAAAATG	GGATGGTGAA	GGGAGAGATT	GAATTACTTA	AGAGATTAAC	AACCATATCA	1800
ATATCAGGAG	TTCCACGGTA	TAATGAAGTG	TACAATAATT	CTAAAAGCCA	TACAGATGAC	1860
CTTAAACCT	ACAATAAAAT	AAGTAATCTT	AATTTGTCTT	CTAATCAGAA	ATCAAAGAAA	1920
TTTGAATTCA	AGTCAACGGA	TATCTACAAT	GATGGATACG	AGACTGTGAG	CTGTTTCCTA	1980
ACAACAGATC	TCAAAAAATA	CTGTCCTTAAT	TGGAGATATG	AATCAACAGC	TCTATTTGGA	2040
GAAACTTGCA	ACCAAATATT	TGGATTAAAT	AAATTGTTTT	ATTGGTTACA	CCCTCGTCTT	2100
GAAGGAAGTA	CAATCTATGT	AGGTGATCCT	TACTGTCCTC	CATCAGATAA	AGAACATATA	2160
TCATTAGAGG	ATCACCTTGA	TTCTGGTTTT	TACGTTTATA	ACCCAAGAGG	GGGTATAGAA	2220
GGATTTTGTG	AAAAATTATG	GACACTCATA	TCTATAAGTG	CAATACATCT	AGCAGCTGTT	2280
AGAATAGGCG	TGAGGGTGAC	TGCAATGGTT	CAAGGAGACA	ATCAAGCTAT	AGCTGTAACC	2340
ACAAGAGTAC	CCAACAATTA	TGACTACAGA	GTTAAGAAGG	AGATAGTTTA	TAAAGATGTA	2400

GTGAGATTTT	TTGATTCAAT	AAGAGAAGTG	ATGGATGATC	TAGGTCATGA	ACTTAAATTA	2460
AATGAAACGA	TTATAAGTAG	CAAGATGTTT	ATATATAGCA	AAAGAATCTA	TTATGATGGG	2520
AGAATTCTTC	CTCAAGCTCT	AAAAGCATT	TCTAGATGTG	TCTTCTGGTC	AGAGACAGTA	2580
ATAGACGAAA	CAAGATCAGC	ATCTTCAAAT	TTGGCAACAT	CATTTGCAAA	AGCAATTGAG	2640
AATGGTTATT	CACCTGTTCT	AGGATATGCA	TGCTCAATTT	TTAAGAACAT	TCAACAACATA	2700
TATATTGCC	TTGGGATGAA	TATCAATCCA	ACTATAACAC	AGAATATCAG	AGATCAGTAT	2760
TTTAGGAATC	CAAATTGGAT	GCAATATGCC	TCTTTAATAC	CTGCTAGTGT	TGGGGGATTC	2820
AATTACATGG	CCATGTCAAG	ATGTTTTGTA	AGGAATATTG	GTGATCCATC	AGTTGCCGCA	2880
TTGGCTGATA	TTAAAAGATT	TATTAAGGCG	AATCTATTAG	ACCGAAGTGT	TCTTTATAGG	2940
ATTATGAATC	AAGAACCAGG	TGAGTCATCT	TTTTTTGGACT	GGGCTTCAGA	TCCATATTCA	3000
TGCAATTTAC	CACAATCTCA	AAATATAACC	ACCATGATAA	AAAATATAAC	AGCAAGGAAT	3060
GTATTACAAG	ATTACACAAA	TCCATTATTA	TCTGGATTAT	TCACAAATAC	AATGATAGAA	3120
GAAGATGAAG	AATTAGCTGA	GTTCTTGATG	GACAGGAAGG	TAATTCTCCC	TAGAGTTGCA	3180
CATGATATTC	TAGATAATTC	TCTCACAGGA	ATTAGAAATG	CCATAGCTGG	AATGTTAGAT	3240
ACGACAAAAT	CACTAATTCG	GGTTGGCATA	AATAGAGGAG	GACTGACATA	TAGTTTGTGTG	3300
AGGAAAATCA	GTAATTACGA	TCTAGTACAA	TATGAAACAC	TAAGTAGGAC	TTTGCGACTA	3360
ATTGTAAGTG	ATAAAATCAA	GTATGAAGAT	ATGTGTTCTGG	TAGACCTTGC	CATAGCATTG	3420
CGACAAAAGA	TGTGGATTCA	TTATCAGGA	GGAAGGATGA	TAAGTGGACT	TGAAACGCCT	3480
GACCCATTAG	AATTACTATC	TGGGGTAGTA	ATAACAGGAT	CAGAACATTG	TAAAATATGT	3540
TATTCTTCAG	ATGGCACAAA	CCCATATACT	TGGATGTATT	TACCCGGTAA	TATCAAAATA	3600
GGATCAGCAG	AAACAGGTAT	ATCGTCATTA	AGAGTTCCTT	ATTTTGGATC	AGTCACTGAT	3660
GAAAGATCTG	AAGCACAATT	AGGATATATC	AAGAATCTTA	GTAAACCTGC	AAAAGCCGCA	3720
ATAAGAATAG	CAATGATATA	TACATGGGCA	TTTGGTAATG	ATGAGATATC	TTGGATGGAA	3780
GCCTCACAGA	TAGCACAAAC	ACGTGCAAAT	TTTACACTAG	ATAGTCTCAA	AATTTTAACA	3840
CCGGTAGCTA	CATCAACAAA	TTTATCACAC	AGATTAAAGG	ATACTGCAAC	TCAGATGAAA	3900
TTCTCCAGTA	CATCATTGAT	CAGAGTCAGC	AGATTCATAA	CAATGTCCAA	TGATAACATG	3960
TCTATCAAAG	AAGCTAATGA	AACCAAAGAT	ACTAATCTTA	TTTATCAACA	AATAATGTTA	4020
ACAGGATTAA	GTGTTTTTCGA	ATATTTATTT	AGATTAAAG	AAACCACAGG	ACACAACCCT	4080
ATAGTTATGC	ATCTGCACAT	AGAAGATGAG	TGTTGTATTA	AAGAAAGTTT	TAATGATGAA	4140
CATATTAATC	CAGAGTCTAC	ATTAGAATTA	ATTCGATATC	CTGAAAGTAA	TGAATTTATT	4200
TATGATAAAG	ACCCACTCAA	AGATGTGGAC	TTATCAAAAC	TTATGGTTAT	TAAAGACCAT	4260
TCTTACACAA	TTGATATGAA	TTATTGGGAT	GATACTGACA	TCATACATGC	AATTTCAATA	4320
TGTACTGCAA	TTACAATAGC	AGATACTATG	TCACAATTAG	ATCGAGATAA	TTTAAAAGAG	4380
ATAATAGTTA	TTGCAAATGA	TGATGATATT	AATAGCTTAA	TCACTGAATT	TTTGACTCTT	4440
GACATACTTG	TATTTCTCAA	GACATTTGGT	GGATTATTAG	TAAATCAATT	TGCATACACT	4500

CTTTATAGTC	TAAAAATAGA	AGGTAGGGAT	CTCATTTGGG	ATTATATAAT	GAGAACACTG	4560
AGAGATACTT	CCCATTCAAT	ATTAAAAGTA	TTATCTAATG	CATTATCTCA	TCCTAAAGTA	4620
TTCAAGAGGT	TCTGGGATTG	TGGAGTTTTA	AACCCTATTT	ATGGTCCTAA	TACTGCTAGT	4680
CAAGACCAGA	TAAAACTTGC	CCTATCTATA	TGTGAATATT	CACTAGATCT	ATTTATGAGA	4740
GAATGGTTGA	ATGGTGTATC	ACTTGAAATA	TACATTTGTG	ACAGCGATAT	GGAAGTTGCA	4800
AATGATAGGA	AACAAGCCTT	TATTTCTAGA	CACCTTTCAT	TTGTTTGTTG	TTTAGCAGAA	4860
ATTGCATCTT	TCGGACCTAA	CCTGTTAAAC	TTAACATACT	TGGAGAGACT	TGATCTATTG	4920
AAACAATATC	TTGAATTAAA	TATTAAAGAA	GACCCTACTC	TTAAATATGT	ACAAATATCT	4980
GGATTATTAA	TTAAATCGTT	CCCATCAACT	GTAACATACG	TAAGAAAGAC	TGCAATCAAA	5040
TATCTAAGGA	TTCGCGGTAT	TAGTCCACCT	GAGGTAATTG	ATGATTGGGA	TCCGGTAGAA	5100
GATGAAAATA	TGCTGGATAA	CATTGTCAAA	ACTATAAATG	ATAACTGTAA	TAAAGATAAT	5160
AAAGGGAATA	AAATTAACAA	TTTCTGGGGA	CTAGCACTTA	AGAACTATCA	AGTCCTTAAA	5220
ATCAGATCTA	TAACAAGTGA	TTCTGATGAT	AATGATAGAC	TAGATGCTAA	TACAAGTGGT	5280
TTGACACTTC	CTCAAGGAGG	GAATTATCTA	TCGCATCAAT	TGAGATTATT	CGGAATCAAC	5340
AGCACTAGTT	GTCTGAAAGC	TCTTGAGTTA	TCACAAATTT	TAATGAAGGA	AGTCAATAAA	5400
GACAAGGACA	GGCTCTTCCT	GGGAGAAGGA	GCAGGAGCTA	TGCTAGCATG	TTATGATGCC	5460
ACATTAGGAC	CTGCAGTTAA	TTATTATAAT	TCAGGTTTGA	ATATAACAGA	TGTAATTGGT	5520
CAACGAGAAT	TGAAAATATT	TCCTTCAGAG	GTATCATTAG	TAGGTAAAAA	ATTAGGAAAT	5580
GTGACACAGA	TTCTTAACAG	GGTAAAAAGTA	CTGTTCAATG	GGAATCCTAA	TTCAACATGG	5640
ATAGGAAATA	TGGAATGTGA	GAGCTTAATA	TGGAGTGAAT	TAAATGATAA	GTCCATTGGA	5700
TTAGTACATT	GTGATATGGA	AGGAGCTATC	GGTAAATCAG	AAGAACTGT	TCTACATGAA	5760
CATTATAGTG	TTATAAGAAT	TACATACTTG	ATTGGGGATG	ATGATGTTGT	TTTAGTTTCC	5820
AAAATTATAC	CTACAATCAC	TCCGAATTGG	TCTAGAATAC	TTTATCTATA	TAAATTATAT	5880
TGGAAAGATG	TAAGTATAAT	ATCACTCAAA	ACTTCTAATC	CTGCATCAAC	AGAATTATAT	5940
CTAATTTCGA	AAGATGCATA	TTGTACTATA	ATGGAACCTA	GTGAAATTGT	TTTATCAAAA	6000
CTTAAAAGAT	TGTCACTCTT	GGAAGAAAAAT	AATCTATTAA	AATGGATCAT	TTTATCAAAG	6060
AAGAGGAATA	ATGAATGGTT	ACATCATGAA	ATCAAAGAAG	GAGAAAGAGA	TTATGGAATC	6120
ATGAGACCAT	ATCATATGGC	ACTACAAATC	TTTGGATTTT	AAATCAATTT	AAATCATCTG	6180
GCGAAAGAAT	TTTTATCAAC	CCCAGATCTG	ACTAATATCA	ACAATATAAT	CCAAAGTTTT	6240
CAGCGAACAA	TAAAGGATGT	TTTATTTGAA	TGGATTAATA	TAACTCATGA	TGATAAGAGA	6300
CATAAATTAG	GCGGAAGATA	TAACATATTC	CCACTGAAAA	ATAAGGGAAA	GTTAAGACTG	6360
CTATCGAGAA	GACTAGTATT	AAGTTGGATT	TCATTATCAT	TATCGACTCG	ATTACTTACA	6420
GGTCGCTTTT	CTGATGAAAA	ATTTGAACAT	AGAGCACAGA	CTGGATATGT	ATCATTAGCT	6480
GATACTGATT	TAGAATCATT	AAAGTTATTG	TCGAAAAACA	TCATTAAGAA	TTACAGAGAG	6540
TGTATAGGAT	CAATATCATA	TTGGTTTCTA	ACCAAAGAAG	TTAAATACT	TATGAAATTG	6600

ATTGGTGGTG	CTAAATTATT	AGGAATTCCC	AGACAATATA	AAGAACCCGA	AGACCAGTTA	6660
TTAGAAAAC	ACAATCAACA	TGATGAATTT	GATATCGATT	AAAACATAAA	TACAATGTCTG	6720
ACGATCCGGC	TGCTAACAAA	GCCCGAAAGG	AAGCTGAGTT	GGCTGCTGCC	ACCGCTGAGC	6780
AATAACTAGC	ATAACCCCTT	GGGGCCTCTA	AACGGGTCTT	GAGGGGTTTT	TTGCTGAAAG	6840
GAGGAACTAT	ATCCGGATCG	AGATCAATTC	TGTGAGCGTA	TGGCAAACGA	AGGAAAAATA	6900
GTTATAGTAG	CCGCACTCGA	TGGGACATTT	CAACGTAAAC	CGTTTAATAA	TATTTTGAAT	6960
CTTATTCCAT	TATCTGAAAT	GGTGGTAAAA	CTAACTGCTG	TGTGTATGAA	ATGCTTTAAG	7020
GAGGCTTCCT	TTTCTAAACG	ATTGGGTGAG	GAAACCGAGA	TAGAAATAAT	AGGAGGTAAT	7080
GATATGTATC	AATCGGTGTG	TAGAAAGTGT	TACATCGACT	CATAATATTA	TATTTTTTAT	7140
CTAAAAAACT	AAAAATAAAC	ATTGATTAAA	TTTTAATATA	ATACTTAAAA	ATGGATGTTG	7200
TGTCGTTAGA	TAAACCGTTT	ATGTATTTTG	AGGAAATTGA	TAATGAGTTA	GATTACGAAC	7260
CAGAAAGTGC	AAATGAGGTC	GCAAAAAAAC	TGCCGTATCA	AGGACAGTTA	AACTATTAC	7320
TAGGAGAATT	ATTTTTTCTT	AGTAAGTTAC	AGCGACACGG	TATATTAGAT	GGTGCCACCG	7380
TAGTGTATAT	AGGATCTGCT	CCCGGTACAC	ATATACGTTA	TTTGAGAGAT	CATTTCTATA	7440
ATTTAGGAGT	GATCATCAAA	TGGATGCTAA	TTGACGGCCG	CCATCATGAT	CCTATTTTAA	7500
ATGGATTGCG	TGATGTGACT	CTAGTGACTC	GGTTCGTTGA	TGAGGAATAT	CTACGATCCA	7560
TCAAAAAACA	ACTGCATCCT	TCTAAGATTA	TTTTAATTTT	TGATGTGAGA	TCCAAACGAG	7620
GAGGAAATGA	ACCTAGTACG	GCGGATTTAC	TAAGTAATTA	CGCTCTACAA	AATGTCATGA	7680
TTAGTATTTT	AAACCCCGTG	GCGTCTAGTC	TTAAATGGAG	ATGCCCGTTT	CCAGATCAAT	7740
GGATCAAGGA	CTTTTATATC	CCACACGGTA	ATAAAATGTT	ACAACCTTTT	GCTCCTTCAT	7800
ATTCAGGGCC	GTCGTTTTAC	AACGTCGTGA	CTGGGAAAC	CCTGGCGTTA	CCCAACTTAA	7860
TCGCCTTGCA	GCACATCCCC	CTTTCGCCAG	CTGGCGTAAT	AGCGAAGAGG	CCCGCACCGA	7920
TCGCCCTTCC	CAACAGTTGC	GCAGCCTGAA	TGGCGAATGG	CGCGACGCGC	CCTGTAGCGG	7980
CGCATTAAGC	GCGGCGGGTG	TGGTGGTTAC	GCGCAGCGTG	ACCGCTACAC	TTGCCAGCGC	8040
CCTAGCGCCC	GCTCCTTTTC	CTTCTTTCCC	TTCTTTTCTC	GCCACGTTTC	CCGGCTTTCC	8100
CCGTCAAGCT	CTAAATCGGG	GGCTCCCTTT	AGGGTTCCGA	TTTAGTGCTT	TACGGCACCT	8160
CGACCCCAAA	AACTTGATT	AGGGTGATGG	TTACGTAGT	GGGCCATCGC	CCTGATAGAC	8220
GGTTTTTCGC	CCTTTGACGT	TGGAGTCCAC	GTTCTTTAAT	AGTGGACTCT	TGTTCCAAAC	8280
TGGAACAACA	CTCAACCCTA	TCTCGGTCTA	TTCTTTTGAT	TTATAAGGGA	TTTTGCCGAT	8340
TTGGGCCTAT	TGGTTAAAAA	ATGAGCTGAT	TTAACAAAAA	TTTAACGCGA	ATTTTAACAA	8400
AATATTAACG	TTTACAATTT	CCCAGGTGGC	ACTTTTCGGG	GAAATGTGCG	CGGAACCCCT	8460
ATTTGTTTAT	TTTTCTAAAT	ACATTCAAAT	ATGTATCCGC	TCATGAGACA	ATAACCCTGA	8520
TAAATGCTTC	AATAATATTG	AAAAAGGAAG	AGTATGAGTA	TTCAACATTT	CCGTGTCGCC	8580
CTTATTCCCT	TTTTTGCGGC	ATTTTGCCCT	CCTGTTTTTG	CTCACCAGAG	AACGCTGGTG	8640
AAAGTAAAG	ATGCTGAAGA	TCAGTTGGGT	GCACGAGTGG	GTTACATCGA	ACTGGATCTC	8700

AACAGCGGTA	AGATCCTTGA	GAGTTTTCGC	CCCGAAGAAC	GTTTTCCAAT	GATGAGCACT	8760
TTTAAAGTTC	TGCTATGTGG	CGCGGTATTA	TCCCGTATTG	ACGCCGGGCA	AGAGCAACTC	8820
GGTCGCCGCA	TACACTATTC	TCAGAATGAC	TTGGTTGAGT	ACTCACCAGT	CACAGAAAAG	8880
CATCTTACGG	ATGGCATGAC	AGTAAGAGAA	TTATGCAGTG	CTGCCATAAC	CATGAGTGAT	8940
AACACTGCGG	CCAAC TTACT	TCTGACAACG	ATCGGAGGAC	CGAAGGAGCT	AACCGCTTTT	9000
TTGCACAACA	TGGGGGATCA	TGTAAC TCGC	CTTGATCGTT	GGGAACCGGA	GCTGAATGAA	9060
GCCATACCAA	ACGACGAGCG	TGACACCACG	ATGCCTGTAG	CAATGGCAAC	AACGTTGCGC	9120
AAACTATTAA	CTGGCGAACT	ACTTACTCTA	GCTTCCCGGC	AACAATTAAT	AGACTGGATG	9180
GAGGCGGATA	AAGTTGCAGG	ACCACTTCTG	CGCTCGGCCC	TTCCGGCTGG	CTGGTTTATT	9240
GCTGATAAAT	CTGGAGCCGG	TGAGCGTGGG	TCTCGCGGTA	TCATTGCAGC	ACTGGGGCCA	9300
GATGGTAAGC	CCTCCCGTAT	CGTAGTTATC	TACACGACGG	GGAGTCAGGC	AACTATGGAT	9360
GAACGAAATA	GACAGATCGC	TGAGATAGGT	GCCTCACTGA	TTAAGCATTG	GTAAC TGTCA	9420
GACCAAGTTT	ACTCATATAT	ACTTTAGATT	GATTTAAAC	TTCATTTTTA	ATTTAAAAGG	9480
ATCTAGGTGA	AGATCCTTTT	TGATAATCTC	ATGACCAAAA	TCCCTTAACG	TGAGTTTTCG	9540
TTCCACTGAG	CGTCAGACCC	CGTAGAAAAG	ATCAAAGGAT	CTTCTTGAGA	TCCTTTTTTT	9600
CTGCGCGTAA	TCTGCTGCTT	GCAAACAAAA	AAACCACCGC	TACCAGCGGT	GGTTTGTTTG	9660
CCGGATCAAG	AGCTACCAAC	TCTTTTTTCCG	AAGGTA ACTG	GCTTCAGCAG	AGCGCAGATA	9720
CCAAATACTG	TCCTTCTAGT	GTAGCCGTAG	TTAGGCCACC	ACTTCAAGAA	CTCTGTAGCA	9780
CCGCCTACAT	ACCTCGCTCT	GCTAATCCTG	TTACCAGTGG	CTGCTGCCAG	TGGCGATAAG	9840
TCGTGTCTTA	CCGGGTTGGA	CTCAAGACGA	TAGTTACCGG	ATAAGGCGCA	GCGGTCGGGC	9900
TGAACGGGGG	GTTTCGTGCAC	ACAGCCCAGC	TTGGAGCGAA	CGACCTACAC	CGAACTGAGA	9960
TACCTACAGC	GTGAGCTATG	AGAAAGCGCC	ACGCTTCCCG	AAGGGAGAAA	GGCGGACAGG	10020
TATCCGGTAA	GCGGCAGGGT	CGGAACAGGA	GAGCGCACGA	GGGAGCTTCC	AGGGGGAAAC	10080
GCCTGGTATC	TTTATAGTCC	TGTCGGGTTT	CGCCACCTCT	GACTTGAGCG	TCGATTTTTG	10140
TGATGCTCGT	CAGGGGGGCG	GAGCCTATGG	AAAAACGCCA	GCAACGCGGC	CTTTTTACGG	10200
TTCTTGCCCT	TTTGCTGGCC	TTTTGCTCAC	ATGTTCTTTC	CTGCGTTATC	CCCTGATTCT	10260
GTGGATAACC	GTATTACCGC	CTTTGAGTGA	GCTGATACCG	CTCGCCGCAG	CCGAACGACC	10320
GAGCGCAGCG	AGTCAGTGAG	CGAGGAAGCG	GAAGAGCGCC	CAATACGCAA	ACCGCCTCTC	10380
CCCGCGCGTT	GGCCGATTCA	TTAATGCAGC	TGGCACGACA	GGTTTCCCGA	CTGGAAAGCG	10440
GGCAGTGAGC	GCAACGCAAT	TAATGTGAGT	TAGCTCACTC	ATTAGGCACC	CCAGGCTTTA	10500
CACTTTTATGC	TTCCGGCTCG	TATGTTGTGT	GGAATTGTGA	GCGGATAACA	ATTTACACAC	10560
GGAAACAGCT	ATGACCATGA	TTACGCCAAG	CTTTTGCGAT	CAATAAATGG	ATCACAACCA	10620
GTATCTCTTA	ACGATGTTCT	TCGCAGATGA	TGATTCATTT	TTTAAGTATT	TGGCTAGTCA	10680
AGATGATGAA	TCTTCATTAT	CTGATATATT	GCAAATCACT	CAATATCTAG	ACTTTCTGTT	10740
ATTATTATTG	ATCCAATCAA	AAAATAAATT	AGAAGCCGTG	GGTCATTGTT	ATGAATCTCT	10800



TTCAGAGGAA TACAGACAAT TGACAAAATT CACAGACTTT CAAGATTTTA AAAAAGTGT 10860  
 TAACAAGGTC CCTATTGTTA CAGATGGAAG GGTCAAACCTT AATAAAGGAT ATTTGTTCTGA 10920  
 CTTTGTGATT AGTTTGATGC GATTCAAAAA AGAATCCTCT CTAGCTACCA CCGCAATAGA 10980  
 TCCTGTTAGA TACATAGATC CTCGTCGCAA TATCGCATTT TCTAACGTGA TGGATATATT 11040  
 AAAGTCGAAT AAAGTGAACA ATAATTAATT CTTTATTGTC ATCATGAACG GCGGACATAT 11100  
 TCAGTTGATA ATCGGCCCCA TGTTTTTCAGG TAAAAGTACA GAATTAATTA GACGAGTTAG 11160  
 ACGTTATCAA ATAGCTCAAT ATAAATGCGT GACTATAAAA TATTCTAACG ATAATAGATA 11220  
 CGGAACGGGA CTATGGACGC ATGATAAGAA TAATTTTGAA GCATTGGAAG CAACTAAACT 11280  
 ATGTGATGTC TTGGAATCAA TTACAGATTT CTCCGTGATA GGTATCGATG AAGGACAGTT 11340  
 CTTTCCAGAC ATTGTTGAAT TGATCTCGAT CCCGCGAAAT TAATACGACT CACTATAGGG 11400  
 AGACCACAAC GGTTCCTCTC TAGCGGGATC AATTCGCCC CTCTCCCTCC CCCCCCCTA 11460  
 ACGTTACTGG CCGAAGCCGC TTGGAATAAG GCCGGTGTGC GTTTGTCTAT ATGTTATTTT 11520  
 CCACCATATT GCCGTCTTTT GGCAATGTGA GGGCCCGGAA ACCTGGCCCT GTCTTCTTGA 11580  
 CGAGCATTCC TAGGGGTCTT TCCCCTCTCG CCAAAGGAAT GCAAGGTCTG TTGAATGTCTG 11640  
 TGAAGGAAGC AGTTCCTCTG GAAGCTTCTT GAAGACAAAC AACGTCTGTA GCGACCCCTT 11700  
 GCAGGCAGCG GAACCCCCCA CCTGGCGACA GGTGCCTCTG CGGCCAAAAG CCACGTGTAT 11760  
 AAGATACACC TGCAAAGGCG GCACAACCCC AGTGCCACGT TGTGAGTTGG ATAGTTGTGG 11820  
 AAAGAGTCAA ATGGCTCTCC TCAAGCGTAT TCAACAAGGG GCTGAAGGAT GCCCAGAAGG 11880  
 TACCCCATTTG TATGGGATCT GATCTGGGGC CTCGGTGCAC ATGCTTTACA TGTGTTTAGT 11940  
 CGAGGTAAAT AAACGTCTAG GCCCCCGAA CCACGGGGAC GTGGTTTTTC TTTGAAAAAC 12000  
 ACGATAATAC C 12011

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GATCGATGCT AGCCC

15

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

15

12

12

12

12

15

12

12

12

12

15

12

12

12

12

15

12

12

12

12

15

12

12

(ii) MOLECULE TYPE: cDNA

TGGTCCTAAT ACTG

14

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 14 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

TGGGCCTAAT ATCG

14

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 30 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

GCATTATCTA GATGTGTCTT CTGGTCAGAG

30

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 30 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

CCTGAATTAT AATAATTAAC TGCAGGTCCT

30

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear



CGCCATGGAA AAATCAGAGA TCCTCTTCT

29

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 31 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CTGGATCCTA ATTGGAGTTG TTACCCATGT A

31

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

AACCATGGCT GAAAAAGGGA AAA

23

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 33 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GGTGAAGCTT AAGATGTGAT TTTACATATT TTA

33

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

AAATAGGATC CCTACAGATC ATTAGATATT AAAAT

35

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:



AACATGCACA	AAGGGCAGGG	TTCTTGGTGT	CTTTATTGTC	AATGGCTTAT	GCCAATCCAG	360
AGCTCTACCT	AACAACAAAT	GGAAGTAATG	CAGATGCCAA	GTATGTCATA	TACATGATTG	420
AGAAAGATCT	AAAACGGCAA	AAGTATGGAG	GATTTGTGGT	TAAGACGAGA	GAGATGATAT	480
ATGAAAAGAC	AACTGATTGG	ATATTTGGAA	GTGACCTGGA	TTATGATCAG	GAAACTATGT	540
TGCAGAACGG	CAGGAACAAT	TCAACAATTG	AAGACCTTGT	CCACACATTT	GGGTATCCAT	600
CATGTTTAGG	AGCTCTTATA	ATACAGATCT	GGATAGTTCT	GGTCAAAGCT	ATCACTAGTA	660
TCTCAGGGTT	AAGAAAAGGC	TTTTTCACCC	GATTGGAAGC	TTTCAGACAA	GATGGAACAG	720
TGCAGGCAGG	GCTGGTATTG	AGCGGTGACA	CAGTGGATCA	GATTGGGTCA	ATCATGCGGT	780
CTCAACAGAG	CTTGGTAACT	CTTATGGTTG	AAACATTAAT	AACAATGAAT	ACCAGCAGAA	840
ATGACCTCAC	AACCATAGAA	AAGAATATAC	AAATTGTTGG	CAACTACATA	AGAGATGCAG	900
GTCTCGCTTC	ATTCTTCAAT	ACAATCAGAT	ATGGAATTGA	GACCAGAATG	GCAGCTTTGA	960
CTCTATCCAC	TCTCAGACCA	GATATCAATA	GATTAAAAGC	TTTGATGGAA	CTGTATTTAT	1020
CAAAGGGACC	ACGCGCTCCT	TTCATCTGTA	TCCTCAGAGA	TCCTATACAT	GGTGAGTTCG	1080
CACCAGGCAA	CTATCCTGCC	ATATGGAGCT	ATGCAATGGG	GGTGGCAGTT	GTACAAAATA	1140
GAGCCATGCA	ACAGTATGTG	ACGGGAAGAT	CATATCTAGA	CATTGATATG	TTCCAGCTAG	1200
GACAAGCAGT	AGCACGTGAT	GCCGAAGCTC	AAATGAGCTC	AACACTGGAA	GATGAACTTG	1260
GAGTGACACA	CGAAGCTAAA	GAAAGCTTGA	AGAGACATAT	AAGGAACATA	AACAGTTCAG	1320
AGACATCTTT	CCACAAACCG	ACAGGTGGAT	CAGCCATAGA	GATGGCAATA	GATGAAGAGC	1380
CAGAACAATT	CGAACATAGA	GCAGATCAAG	AACAAAATGG	AGAACCTCAA	TCATCCATAA	1440
TTCAATATGC	CTGGGCAGAA	GGAAATAGAA	GCGATGATCA	GACTGAGCAA	GCTACAGAAT	1500
CTGACAATAT	CAAGACCGAA	CAACAAAACA	TCAGAGACAG	ACTAAACAAG	AGACTCAACG	1560
ACAAGAAGAA	ACAAAGCAGT	CAACCACCCA	CTAATCCCAC	AAACAGAACA	AACCAGGACG	1620
AAATAGATGA	TCTGTTTAAC	GCATTTGGAA	GCAACTAATC	GAATCAACAT	TTTAATCTAA	1680
ATCAATAATA	AATAAGAAAA	ACTTAGGATT	AAAGAATCCT	ATCATACCGG	AATATAGGGT	1740
GGTAAATTTA	GAGTCTGCTT	GAAACTCAAT	CAATAGAGAG	TTGATGGAAA	GCGATGCTAA	1800
AAACTATCAA	ATCATGGATT	CTTGGAAGA	GGAATCAAGA	GATAAATCAA	CTAATATCTC	1860
CTCGGCCCTC	AACATCATTG	AATTCATACT	CAGCACCGAC	CCCCAAGAAG	ACTTATCGGA	1920
AAACGACACA	ATCAACACAA	GAACCCAGCA	ACTCAGTGCC	ACCATCTGTC	AACCAGAAAT	1980
CAAACCAACA	GAAACAAGTG	AGAAAGATAG	TGGATCAACT	GACAAAAATA	GACAGTCTGG	2040
GTCATCACAC	GAATGTACAA	CAGAAGCAAA	AGATAGAAAC	ATTGATCAGG	AAACTGTACA	2100
GAGAGGACCT	GGGAGAAGAA	GCAGCTCAGA	TAGTAGAGCT	GAGACTGTGG	TCTCTGGAGG	2160
AATCCCCAGA	AGCATCACAG	ATTCTAAAAA	TGGAACCCAA	AACACGGAGG	ATATTGATCT	2220
CAATGAAATT	AGAAAGATGG	ATAAGGACTC	TATTGAGGGG	AAAATGCGAC	AATCTGCAAA	2280
TGTTCCAAGC	GAGATATCAG	GAAGTGATGA	CATATTTACA	ACAGAACAAA	GTAGAAACAG	2340
TGATCATGGA	AGAAGCCTGG	AATCTATCAG	TACACCTGAT	ACAAGATCAA	TAAGTGTGTG	2400

TACTGCTGCA	ACACCAGATG	ATGAAGAAGA	AATACTAATG	AAAAATAGTA	GGACAAAGAA	2460
AAGTTCTTCA	ACACATCAAG	AAGATGACAA	AAGAATTAAA	AAAGGGGGAA	AAGGGAAAAG	2520
CTGGTTTAA	AAATCAAAA	ATACCGACAA	CCAGATACCA	ACATCAGACT	ACAGATCCAC	2580
ATCAAAAGGG	CAGAAGAAAA	TCTCAAAGAC	AACAACCACC	AACACCGACA	CAAAGGGGCA	2640
AACAGAAATA	CAGACAGAAT	CATCAGAAAC	ACAATCCTCA	TCATGGAATC	TCATCATCGA	2700
CAACAACACC	GACCGGAACG	AACAGACAAG	CACAACCTCT	CCAACAACAA	CTTCCAGATC	2760
AACTTATACA	AAAGAATCGA	TCCGAACAAA	CTCTGAATCC	AAACCCAAGA	CACAAAAGAC	2820
AAATGGAAAG	GAAAGGAAGG	ATACAGAAGA	GAGCAATCGA	TTTACAGAGA	GGGCAATTAC	2880
TCTATTGCAG	AATCTTGGTG	TAATTCAATC	CACATCAAAA	CTAGATTTAT	ATCAAGACAA	2940
ACGAGTTGTA	TGTGTAGCAA	ATGTACTAAA	CAATGTAGAT	ACTGCATCAA	AGATAGATTT	3000
CCTGGCAGGA	TTAGTCATAG	GGGTTTCAAT	GGACAACGAC	ACAAAATTAA	CACAGATACA	3060
AAATGAAATG	CTAAACCTCA	AAGCAGATCT	AAAGAAAATG	GACGAATCAC	ATAGAAGATT	3120
GATAGAAAAT	CAAAGAGAAC	AACTGTCATT	GATCACGTCA	CTAATTTCAA	ATCTCAAAAT	3180
TATGACTGAG	AGAGGAGGAA	AGAAAAGACCA	AAATGAATCC	AATGAGAGAG	TATCCATGAT	3240
CAAAACAAAA	TTGAAAGAAG	AAAAGATCAA	GAAGACCAGG	TTTGACCCAC	TTATGGAGGC	3300
ACAAGGCATT	GACAAGAATA	TACCCGATCT	ATATCGACAT	GCAGGAGATA	CACTAGAGAA	3360
CGATGTACAA	GTTAAATCAG	AGATATTAAG	TTCATACAAT	GAGTCAAATG	CAACAAGACT	3420
AATACCCAAA	AAAGTGAGCA	GTACAATGAG	ATCACTAGTT	GCAGTCATCA	ACAACAGCAA	3480
TCTCTCACAA	AGCACAAAAC	AATCATACAT	AAACGAACTC	AAACGTTGCA	AAAATGATGA	3540
AGAAGTATCT	GAATTAATGG	ACATGTTCAA	TGAAGATGTC	AACAATTGCC	AATGATCCAA	3600
CAAAGAAACG	ACACCGAACA	AACAGACAAG	AAACAACAGT	AGATCAAAAC	CTGTCAACAC	3660
ACACAAAATC	AAGCAGAATG	AAACAACAGA	TATCAATCAA	TATACAAATA	AGAAAAACTT	3720
AGGATTAAAG	AATAAATTAA	TCCTTGTTCCA	AAATGAGTAT	AACTAACTCT	GCAATATACA	3780
CATTCCCAGA	ATCATCATTC	TCTGAAAATG	GTCATATAGA	ACCATTACCA	CTCAAAGTCA	3840
ATGAACAGAG	GAAAGCAGTA	CCCCACATTA	GAGTTGCCAA	GATCGGAAAT	CCACCAAAAC	3900
ACGGATCCCG	GTATTTAGAT	GTCTTCTTAC	TCGGCTTCTT	CGAGATGGAA	CGAATCAAAG	3960
ACAAATACGG	GAGTGTGAAT	GATCTCGACA	GTGACCCGAG	TTACAAAGTT	TGTGGCTCTG	4020
GATCATTACC	AATCGGATTG	GCTAAGTACA	CTGGGAATGA	CCAGGAATTG	TTACAAGCCG	4080
CAACCAAAC	GGATATAGAA	GTGAGAAGAA	CAGTCAAAGC	GAAAGAGATG	GTTGTTTACA	4140
CGGTACAAAA	TATAAAACCA	GAAGTGTACC	CATGGTCCAA	TAGACTAAGA	AAAGGAATGC	4200
TGTTTCGATG	CAACAAAGTT	GCTCTTGCTC	CTCAATGTCT	TCCACTAGAT	AGGAGCATAA	4260
AATTTAGAGT	AATCTTCGTG	AATTGTACGG	CAATTGGATC	AATAACCTTG	TTCAAAATTC	4320
CTAAGTCAAT	GGCATCACTA	TCGTTAACCA	ACACAATATC	AATCAATCTG	CAGGTACACA	4380
TAAAAACAGG	GGTTCAGACT	GATTCTAAAG	GGATAGTTCA	AATTTTGGAT	GAGAAAGGCG	4440
AAAAATCACT	GAATTTTCATG	GTCCATCTCG	GATTGATCAA	AAGAAAAGTA	GGCAGAATGT	4500



ACTCTGTTGA	ATACTGTAAA	CAGAAAAATCG	AGAAAAATGAG	ATTGATATTT	TCTTTAGGAC	4560
TAGTTGGAGG	AATCAGTCTT	CATGTCAATG	CAACTGGGTC	CATATCAAAA	ACACTAGCAA	4620
GTCAGCTGGT	ATTCAAAAGA	GAGATTTGTT	ATCCTTTAAT	GGATCTAAAT	CCGCATCTCA	4680
ATCTAGTTAT	CTGGGCTTCA	TCAGTAGAGA	TTACAAGAGT	GGATGCAATT	TTCCAACCTT	4740
CTTTACCTGG	CGAGTTCAGA	TACTATCCTA	ATATTATTGC	AAAAGGAGTT	GGGAAAATCA	4800
AACAATGGAA	CTAGTAATCT	CTATTTTAGT	CCGGACGTAT	CTATTAAGCC	GAAGCAAATA	4860
AAGGATAATC	AAAAACTTAG	GACAAAAGAG	GTCAATACCA	ACAAC TATTA	GCAGTCACAC	4920
TCGCAAGAAT	AAGAGAGAAG	GGACCAAAAA	AGTCAAATAG	GAGAAATCAA	AACAAAAGGT	4980
ACAGAACACC	AGAACAACAA	AATCAAAACA	TCCAAC TCAC	TCAAAACAAA	AATTCCAAAA	5040
GAGACCGGCA	ACACAACAAG	CACTGAACAC	AATGCCAACT	TCAATACTGC	TAATTATTAC	5100
AACCATGATC	ATGGCATCTT	TCTGCCAAAT	AGATATCACA	AAACTACAGC	ACGTAGGTGT	5160
ATTGGTCAAC	AGTCCCAAAG	GGATGAAGAT	ATCACAAAAC	TTTGAAACAA	GATATCTAAT	5220
TTTGAGCCTC	ATACCAAAAA	TAGAAGACTC	TAACTCTTGT	GGTGACCAAC	AGATCAAGCA	5280
ATACAAGAAG	TTATTGGATA	GACTGATCAT	CCCTTTATAT	GATGGATTAA	GATTACAGAA	5340
AGATGTGATA	GTAACCAATC	AAGAATCCAA	TGAAAACACT	GATCCCAGAA	CAAAACGATT	5400
CTTTGGAGGG	GTAATTGGAA	CCATTGCTCT	GGGAGTAGCA	ACCTCAGCAC	AAATTACAGC	5460
GGCAGTTGCT	CTGGTTGAAG	CCAAGCAGGC	AAGATCAGAC	ATCGAAAAAC	TCAAAGAAGC	5520
AATTAGGGAC	ACAAACAAAG	CAGTGCAGTC	AGTTCAGAGC	TCCATAGGAA	ATTTAATAGT	5580
AGCAATTAAA	TCAGTCCAGG	ATTATGTTAA	CAAAGAAATC	GTGCCATCGA	TTGCGAGGCT	5640
AGGTTGTGAA	GCAGCAGGAC	TTCAATTAGG	AATTGCATTA	ACACAGCATT	ACTCAGAATT	5700
AACAAACATA	TTTGGTGATA	ACATAGGATC	GTTACAAGAA	AAAGGAATAA	AATTACAAGG	5760
TATAGCATCA	TTATACCGCA	CAAATATCAC	AGAAATATTC	ACAACATCAA	CAGTTGATAA	5820
ATATGATATC	TATGATCTGT	TATTTACAGA	ATCAATAAAG	GTGAGAGTTA	TAGATGTTGA	5880
CTTGAATGAT	TACTCAATCA	CCCTCCAAGT	CAGACTCCCT	TTATTAACTA	GGCTGCTGAA	5940
CACTCAGATC	TACAAAGTAG	ATTCCATATC	ATATAACATC	CAAAACAGAG	AATGGTATAT	6000
CCCTCTTCCC	AGCCATATCA	TGACGAAAGG	GGCATTCTTA	GGTGAGAGCAG	ACGTCAAAGA	6060
ATGTATAGAA	GCATT CAGCA	GCTATATATG	CCCTTCTGAT	CCAGGATTTG	TATTAAACCA	6120
TGAAATAGAG	AGCTGCTTAT	CAGGAAACAT	ATCCCAATGT	CCAAGAACAA	CGGTCACATC	6180
AGACATTGTT	CCAAGATATG	CATTTGTCAA	TGGAGGAGTG	GTTGCAAAC T	GTATAACAAC	6240
CACCTGTACA	TGCAACGGAA	TTGGTAATAG	AATCAATCAA	CCACCTGATC	AAGGAGTAAA	6300
AATTATAACA	CATAAAGAAT	GTAGTACTGT	GGGTATCAAC	GGAATGCTGT	TCAATACAAA	6360
TAAAGAAGGA	ACTCTTG CAT	TCTATACACC	AAATGATATA	ACACTAAACA	ATTCTGTTAC	6420
ACTGGATCCA	ATTGACATAT	CAATCGAGCT	CAACAAGGCC	AAATCAGATC	TAGAAGAATC	6480
AAAAGAATGG	ATAAGAAGGT	CAAATCAAAA	ACTAGATTCT	ATTGGAAAT T	GGCATCAATC	6540
TAGCACTACA	ATCATAATTA	TTTTGATAAT	GATCATTATA	TTGTTTATAA	TTAATATAAC	6600

GATAATTACA	ATTGCAATTA	AGTATTACAG	AATTCAAAAG	AGAAATCGAG	TGGATCAAAA	6660
TGACAAGCCA	TATGTACTAA	CAAACAAATA	ACATATCTAC	AGATCATTAG	ATATTA AAAAT	6720
TATAAAAAAC	TTAGGAGTAA	AGTTACGCAA	TCCAAC TCTA	CTCATATAAT	TGAGGAAGGA	6780
CCCAATAGAC	AAATCCAAAT	TCGAGATGGA	ATACTGGAAG	CATACCAATC	ACGGAAAGGA	6840
TGCTGGTAAT	GAGCTGGAGA	CGTCTATGGC	TACTCATGGC	AACAAGCTCA	CTAATAAGAT	6900
AATATACATA	TTATGGACAA	TAATCCTGGT	GTTATTATCA	ATAGTCTTCA	TCATAGTGCT	6960
AATTAATTCC	ATCAAAAGTG	AAAAGGCCCA	CGAATCATTG	CTGCAAGACA	TAAATAATGA	7020
GTTTATGGAA	ATTACAGAAA	AGATCCAAAT	GGCATCGGAT	AATACCAATG	ATCTAATACA	7080
GTCAGGAGTG	AATACAAGGC	TTCTTACAAT	TCAGAGTCAT	GTCCAGAATT	ACATACCAAT	7140
ATCATTGACA	CAACAGATGT	CAGATCTTAG	GAAATTCATT	AGTGAAATTA	CAATTAGAAA	7200
TGATAATCAA	GAAGTGCTGC	CACAAAGAAT	AACACATGAT	GTAGGTATAA	AACCTTTAAA	7260
TCCAGATGAT	TTTTGGAGAT	GCACGTCTGG	TCTTCCATCT	TTAATGAAAA	CTCCAAAAAT	7320
AAGGTTAATG	CCAGGGCCGG	GATTATTAGC	TATGCCAACG	ACTGTTGATG	GCTGTGTTAG	7380
AACTCCGTCT	TTAGTTATAA	ATGATCTGAT	TTATGCTTAT	ACCTCAAATC	TAATTACTCG	7440
AGGTTGTCAG	GATATAGGAA	AATCATATCA	AGTCTTACAG	ATAGGGATAA	TAACTGTAAA	7500
CTCAGACTTG	GTACCTGACT	TAAATCCTAG	GATCTCTCAT	ACCTTTAACA	TAAATGACAA	7560
TAGGAAGTCA	TGTTCTCTAG	CACTCCTAAA	TATAGATGTA	TATCAACTGT	GTTCAACTCC	7620
CAAAGTTGAT	GAAAGATCAG	ATTATGCATC	ATCAGGCATA	GAAGATATTG	TACTTGATAT	7680
TGTCAATTAT	GATGGTTCAA	TCTCAACAAC	AAGATTTAAG	AATAATAACA	TAAGCTTTGA	7740
TCAACCATAT	GCTGCACTAT	ACCCATCTGT	TGGACCAGGG	ATATACTACA	AAGGCAAAAT	7800
AATATTTCTC	GGGTATGGAG	GTCTTGAACA	TCCAATAAAT	GAGAATGTAA	TCTGCAACAC	7860
AACTGGGTGC	CCCGGGAAAA	CACAGAGAGA	CTGTAATCAA	GCATCTCATA	GTA CTTG GTT	7920
TTCAGATAGG	AGGATGGTCA	ACTCCATCAT	TGTGGCTGAC	AAAGGCTTAA	ACTCAATTCC	7980
AAAATTGAAA	GTATGGACGA	TATCTATGCG	ACAAAATTAC	TGGGGGTCAG	AAGGAAGGTT	8040
ACTTCTACTA	GGTAACAAGA	TCTATATATA	TACAAGATCT	ACAAGTTGGC	ATAGCAAGTT	8100
ACAATTAGGA	ATAATTGATA	TTACTGATTA	CAGTGATATA	AGGATAAAAT	GGACATGGCA	8160
TAATGTGCTA	TCAAGACCAG	GAAACAATGA	ATGTCCATGG	GGACATTCAT	GTCCAGATGG	8220
ATGTATAACA	GGAGTATATA	CTGATGCATA	TCCACTCAAT	CCCACAGGGA	GCATTGTGTC	8280
ATCTGTCATA	TTAGACTCAC	AAAAATCGAG	AGTGAACCCA	GTCATAACTT	ACTCAACAGC	8340
AACCGAAAGA	GTAAACGAGC	TGGCCATCCT	AAACAGAACA	CTCTCAGCTG	GATATACAAC	8400
AACAAGCTGC	ATTACACACT	ATAACAAAGG	ATATTGTTTT	CATATAGTAG	AAATAAATCA	8460
TAAAAGCTTA	AACACATTTC	AACCCATGTT	GTTCAAAACA	GAGATTCCAA	AAAGCTGCAG	8520
TTAATCATAA	TTAACCATAA	TATGCATCAA	TCTATCTATA	ATACAAGTAT	ATGATAAGTA	8580
ATCAGCAATC	AGACAATAGA	CAAAAGGGAA	ATATAAAAAA	CTTAGGAGCA	AAGCGTGCTC	8640
GGGAAATGGA	CACTGAATCT	AACAATGGCA	CTGTATCTGA	CATACTCTAT	CCTGAGTGTC	8700

ACCTTAACTC	TCCTATCGTT	AAAGGTAAAA	TAGCACAAATT	ACACACTATT	ATGAGTCTAC	8760
CTCAGCCTTA	TGATATGGAT	GACGACTCAA	TACTAGTTAT	CACTAGACAG	AAAATAAAAC	8820
TTAATAAATT	GGATAAAAGA	CAACGATCTA	TTAGAAGATT	AAAATTAATA	TTAACTGAAA	8880
AAGTGAATGA	CTTAGGAAAA	TACACATTTA	TCAGATATCC	AGAAATGTCA	AAAGAAATGT	8940
TCAAATTATA	TATACCTGGT	ATTAACAGTA	AAGTGA CTGA	ATTATTACTT	AAAGCAGATA	9000
GAACATATAG	TCAAATGACT	GATGGATTAA	GAGATCTATG	GATTAATGTG	CTATCAAAAT	9060
TAGCCTCAAA	AAATGATGGA	AGCAATTATG	ATCTTAATGA	AGAAATTAAT	AATATATCGA	9120
AAGTTCACAC	AACCTATAAA	TCAGATAAAT	GGTATAATCC	ATTCAAAACA	TGGTTTACTA	9180
TCAAGTATGA	TATGAGAAGA	TTACAAAAAG	CTCGAAATGA	GATCACTTTT	AATGTTGGGA	9240
AGGATTATAA	CTTGTTAGAA	GACCAGAAGA	ATTTCTTATT	GATACATCCA	GAATTGGTTT	9300
TGATATTAGA	TAAACAAAAC	TATAATGGTT	ATCTAATTAC	TCCTGAATTA	GTATTGATGT	9360
ATTGTGACGT	AGTCGAAGGC	CGATGGAATA	TAAGTGCATG	TGCTAAGTTA	GATCCAAAAT	9420
TACAATCTAT	GTATCAGAAA	GGTAATAACC	TGTGGGAAGT	GATAGATAAA	TTGTTTCCAA	9480
TTATGGGAGA	AAAGACATTT	GATGTGATAT	CGTTATTAGA	ACCACTTGCA	TTATCCTTAA	9540
TTCAA CTCA	TGATCCTGTT	AAACA ACTAA	GAGGAGCTTT	TTTAAATCAT	GTGTTATCCG	9600
AGATGGAATT	AATATTTGAA	TCTAGAGAAT	CGATTAAGGA	ATTTCTGAGT	G TAGATTACA	9660
TTGATAAAAT	TTTAGATATA	TTTAATAAGT	CTACAATAGA	TGAAATAGCA	GAGATTTTCT	9720
CTTTTTTTTAG	AACATTTGGG	CATCCTCCAT	TAGAAGCTAG	TATTGCAGCA	GAAAAGGTTA	9780
GAAAATATAT	GTATATTGGA	AAACAATTAA	AATTTGACAC	TATTAATAAA	TGTCATGCTA	9840
TCTTCTGTAC	AATAATAATT	AACGGATATA	GAGAGAGGCA	TGGTGGACAG	TGGCCTCCTG	9900
TGACATTACC	TGATCATGCA	CACGAATTCA	TCATAAATGC	TTACGGTTCA	AACTCTGCGA	9960
TATCATATGA	AAATGCTGTT	GATTATTACC	AGAGCTTTAT	AGGAATAAAA	TTCAATAAAT	10020
TCATAGAGCC	TCAGTTAGAT	GAGGATTTGA	CAATTTATAT	GAAAGATAAA	GCATTATCTC	10080
CAAAAAAATC	AAATTGGGAC	ACAGTTTATC	CTGCATCTAA	TTTACTGTAC	CGTACTAACG	10140
CATCCAACGA	ATCACGAAGA	TTAGTTGAAG	TATTTATAGC	AGATAGTAAA	TTTGATCCTC	10200
ATCAGATATT	GGATTATGTA	GAATCTGGGG	ACTGGTTAGA	TGATCCAGAA	TTTAATATTT	10260
CTTATAGTCT	TAAAGAAAAA	GAGATCAAAC	AGGAAGGTAG	ACTCTTTGCA	AAAATGACAT	10320
ACAAAATGAG	AGCTACACAA	GTTTTATCAG	AGACCCTACT	TGCAAATAAC	ATAGGAAAAT	10380
TCTTTCAAGA	AAATGGGATG	GTGAAGGGAG	AGATTGAATT	ACTTAAGAGA	TTAACAACCA	10440
TATCAATATC	AGGAGTTCCA	CGGTATAATG	AAGTGTACAA	TAATTCTAAA	AGCCATACAG	10500
ATGACCTTAA	AACCTACAAT	AAAATAAGTA	ATCTTAATTT	GTCTTCTAAT	CAGAAATCAA	10560
AGAAATTTGA	ATTCAAAGTCA	ACGGATATCT	ACAATGATGG	ATACGAGACT	GTGAGCTGTT	10620
TCCTAACAAC	AGATCTCAAA	AAATACTGTC	TTAATTGGAG	ATATGAATCA	ACAGCTCTAT	10680
TTGGAGAAAC	TTGCAACCAA	ATATTTGGAT	TAAATAAATT	GTTTAATTGG	TTACACCCTC	10740
GTCTTG AAGG	AAGTACAATC	TATGTAGGTG	ATCCTTACTG	TCCTCCATCA	GATAAAGAAC	10800

ATATATCATT	AGAGGATCAC	CCTGATTCTG	GTTTTTACGT	TCATAACCCA	AGAGGGGGTA	10860
TAGAAGGATT	TTGTCAAAAA	TTATGGACAC	TCATATCTAT	AAGTGCAATA	CATCTAGCAG	10920
CTGTTAGAAT	AGGCGTGAGG	GTGACTGCAA	TGGTTCAAGG	AGACAATCAA	GCTATAGCTG	10980
TAACCACAAG	AGTACCCAAC	AATTATGACT	ACAGAGTTAA	GAAGGAGATA	GTTTATAAAG	11040
ATGTAGTGAG	ATTTTTTGAT	TCATTAAGAG	AAGTGATGGA	TGATCTAGGT	CATGAACCTA	11100
AATTAAATGA	AACGATTATA	AGTAGCAAGA	TGTTCAATATA	TAGCAAAAGA	ATCTATTATG	11160
ATGGGAGAAT	TCTTCCTCAA	GCTCTAAAAG	CATTATCTAG	ATGTGTCTTC	TGGTCAGAGA	11220
CAGTAATAGA	CGAAACAAGA	TCAGCATCTT	CAAATTTGGC	AACATCATTT	GCAAAAGCAA	11280
TTGAGAATGG	TTATTCACCT	GTTCTAGGAT	ATGCATGCTC	AATTTTTAAG	AATATTCAAC	11340
AACTATATAT	TGCCCTTGGG	ATGAATATCA	ATCCAACCTAT	AACACAGAAT	ATCAGAGATC	11400
AGTATTTTAG	GAATCCAAAT	TGGATGCAAT	ATGCCTCTTT	AATACCTGCT	AGTGTGGGG	11460
GATTCAATCA	CATGGCGATG	TCAAGATGTT	TTGTAAGGAA	TATTGGTGAT	CCATCAGTTG	11520
CCGCATTGGC	TGATATTAAA	AGATTTATTA	AGGCGAATCT	ATTAGACCGA	AGTGTCTTTT	11580
ATAGGATTAT	GAATCAAGAA	CCAGGTGAGT	CATCTTTTTT	TGATTGGGCT	TCAGATCCAT	11640
ATTCATGCAA	TTTACCACAA	TCTCAAAATA	TAACCACCAT	GATAAAAAAT	ATAACAGCAA	11700
GGAATGTATT	ACAAGATTCA	CCAAATCCAT	TATTATCTGG	ATTATTCACA	AATACAATGA	11760
TAGAAGAAGA	TGAAGAATTA	GCTGAGTTCC	TGATGGACAG	GAAGGTAATT	CTCCCTAGAG	11820
TTGCACATGA	TATTCTAGAT	AATTCTCTCA	CAGGAATTAG	AAATGCCATA	GCTGGAATGT	11880
TAGATACGAC	AAAATCACTA	ATTCGGGTTG	GCATAAATAG	AGGAGGACTG	ACATATAGTT	11940
TGTTGAGGAA	AATCAGTAAT	TACGATCTAG	TACAATATGA	AACACTAAGT	AGGACTTTGC	12000
GACTAATTGT	AAGTGATAAA	ATCAAGTATG	AAGATATGTG	TTCGGTAGAC	CTTGCCATAG	12060
CATTGCGACA	AAAGATGTGG	ATTCATTTAT	CAGGAGGAAG	GATGATAAGT	GGACTTGAAA	12120
CGCCTGACCC	ATTAGAATTA	CTATCTGGGG	TAGTAATAAC	AGGATCAGAA	CATTGTAAAA	12180
TATGTTATTC	TTCAGATGGC	ACAAACCCAT	ATACTTGGAT	GTATTTACCC	GGTAATATCA	12240
AAATAGGATC	AGCAGAAACA	GGTATATCGT	CATTAAGAGT	TCCTTATTTT	GGATCAGTCA	12300
CTGATGAAAG	ATCTGAAGCA	CAATTAGGAT	ATATCAAGAA	TCTTAGTAAA	CCTGCAAAAG	12360
CCGCAATAAG	AATAGCAATG	ATATATACAT	GGGCATTTGG	TAATGATGAG	ATATCTTGGA	12420
TGGAAGCCTC	ACAGATAGCA	CAAACACGTG	CAAATTTTAC	ACTAGATAGT	CTCAAAATTT	12480
TAACACCGGT	AGCTACATCA	ACAAATTTAT	CACACAGATT	AAAGGATACT	GCAACTCAGA	12540
TGAAATTCTC	CAGTACATCA	TTGATCAGAG	TCAGCAGATT	CATAACAATG	TCCAATGATA	12600
ACATGTCTAT	CAAAGAAGCT	AATGAAACCA	AAGATACTAA	TCTTATTTAT	CAACAAATAA	12660
TGTTAACAGG	ATTAAGTGTT	TTCGAATATT	TATTTAGATT	AAAAGAAACC	ACAGGACACA	12720
ACCCTATAGT	TATGCATCTG	CACATAGAAG	ATGAGTGTTG	TATTAAAGAA	AGTTTAAATG	12780
ATGAACATAT	TAATCCAGAG	TCTACATTAG	AATTAATTCG	ATATCCTGAA	AGTAATGAAT	12840
TTATTTATGA	TAAAGACCCA	CTCAAAGATG	TGGACTTATC	AAAACCTTATG	GTTATTAAAG	12900

ACCATTCTTA	CACAATTGAT	ATGAATTATT	GGGATGATAC	TGACATCAT	CATGCAATTT	12960
CAATATGTAC	TGCAATTACA	ATAGCAGATA	CTATGTCACA	ATTAGATCGA	GATAATTTAA	13020
AAGAGATAAT	AGTTATTGCA	AATGATGATG	ATATTAATAG	CTTAATCACT	GAATTTTTGA	13080
CTCTTGACAT	ACTTGTATTT	CTCAAGACAT	TTGGTGGATT	ATTAGTAAAT	CAATTTGCAT	13140
ACACTCTTTA	TAGTCTAAAA	ATAGAAGGTA	GGGATCTCAT	TTGGGATTAT	ATAATGAGAA	13200
CACTGAGAGA	TACTTCCCAT	TCAATATTAA	AAGTATTATC	TAATGCATTA	TCTCATCCTA	13260
AAGTATTCAA	GAGGTTCTGG	GATTGTGGAG	TTTTAAACCC	TATTTATGGG	CCTAATATCG	13320
CTAGTCAAGA	CCAGATAAAA	CTTGCCCTAT	CTATATGTGA	ATATTCACTA	GATCTATTTA	13380
TGAGAGAATG	GTTGAATGGT	GTATCACTTG	AAATATACAT	TTGTGACAGC	GATATGGAAG	13440
TTGCAAATGA	TAGGAAACAA	GCCTTTATTT	CTAGACACCT	TTCATTTGTT	TGTTGTTTAG	13500
CAGAAATTGC	ATCTTTCGGA	CCTAACCTGT	TAAACTTAAC	ATACTTGGAG	AGACTTGATC	13560
TATTGAAACA	ATATCTTGAA	TTAAATATTA	AAGAAGACCC	TACTCTTAAA	TATGTACAAA	13620
TATCTGGATT	ATTAATTTAA	TCGTTCCCAT	CAACTGTAAC	ATACGTAAGA	AAGACTGCAA	13680
TCAAATATCT	AAGGATTTCG	GGTATTAGTC	CACCTGAGGT	AATTGATGAT	TGGGATCCGG	13740
TAGAAGATGA	AAATATGCTG	GATAACATTG	TCAAACTAT	AAATGATAAC	TGTAATAAAG	13800
ATAATAAAGG	GAATAAAATT	AACAATTTCT	GGGGACTAGC	ACTTAAGAAC	TATCAAGTCC	13860
TTAAAATCAG	ATCTATAACA	AGTGATTCTG	ATGATAATGA	TAGACTAGAT	GCTAATACAA	13920
TTGGTTTTGAC	ACTTCCTCAA	GGAGGGAATT	ATCTATCGCA	TCAATTGAGA	TTATTCGGAA	13980
TCAACAGCAC	TAGTTGTCTG	AAAGCTCTTG	AGTTATCACA	AATTTTAATG	AAGGAAGTCA	14040
ATAAAGACAA	GGACAGGCTC	TTCTTGGGAG	AAGGAGCAGG	AGCTATGCTA	GCATGTTATG	14100
ATGCCACATT	AGGACCTGCA	GTTAATTATT	ATAATTCAGG	TTTGAATATA	ACAGATGTAA	14160
TTGGTCAACG	AGAATTGAAA	ATATTTCCTT	CAGAGGTATC	ATTAGTAGGT	AAAAAATTAG	14220
GAAATGTGAC	ACAGATTCTT	AACAGGGTAA	AAGTACTGTT	CAATGGGAAT	CCTAATTCAA	14280
CATGGATAGG	AAATATGGAA	TGTGAGAGCT	TAATATGGAG	TGAATTAAAT	GATAAGTCCA	14340
TTGGATTAGT	ACATTGTGAT	ATGGAAGGAG	CTATCGGTAA	ATCAGAAGAA	ACTGTTCTAC	14400
ATGAACATTA	TAGTGTTATA	AGAATTACAT	ACTTGATTGG	GGATGATGAT	GTTGTTTTAG	14460
TTTCCAAAT	TATACCTACA	ATCACTCCGA	ATTGGTCTAG	AATACTTTAT	CTATATAAAT	14520
TATATTGGAA	AGATGTAAGT	ATAATATCAC	TCAAACTTC	TAATCCTGCA	TCAACAGAAT	14580
TATATCTAAT	TTCGAAAGAT	GCATATTGTA	CTATAATGGA	ACCTAGTGAA	ATTGTTTTAT	14640
CAAAACTTAA	AAGATTGTCA	CTCTTGGAAG	AAAATAATCT	ATTAAAATGG	ATCATTTTAT	14700
CAAAGAAGAG	GAATAATGAA	TGGTTACATC	ATGAAATCAA	AGAAGGAGAA	AGAGATTATG	14760
GAATCATGAG	ACCATATCAT	ATGGCACTAC	AAATCTTTGG	ATTTCAAATC	AATTTAAATC	14820
ATCTGGCGAA	AGAATTTTTA	TCAACCCAG	ATCTGACTAA	TATCAACAAT	ATAATCCAAA	14880
GTTTTTCAGCG	AACAATAAAG	GATGTTTTAT	TTGAATGGAT	TAATATAACT	CATGATGATA	14940
AGAGACATAA	ATTAGGCGGA	AGATATAACA	TATTTCCACT	GAAAAATAAG	GGAAAAGTTAA	15000

GACTGCTATC GAGAAGACTA GTATTAAGTT GGATTTTCATT ATCATTATCG ACTCGATTAC 15060  
 TTACAGGTCG CTTTCCTGAT GAAAAATTTG AACATAGAGC ACAGACTGGA TATGTATCAT 15120  
 TAGCTGATAC TGATTTAGAA TCATTAAAGT TATTGTCGAA AAACATCATT AAGAATTACA 15180  
 GAGAGTGTAT AGGATCAATA TCATATTGGT TTCTAACCAA AGAAGTTAAA ATACTTATGA 15240  
 AATTGATCGG TGGTGCTAAA TTATTAGGAA TTCCCAGACA ATATAAAGAA CCCGAAGACC 15300  
 AGTTATTAGA AAACATAAAT CAACATGATG AATTTGATAT CGATTAAAAC ATAAATACAA 15360  
 TGAAGATATA TCCTAACCTT TATCTTTAAG CCTAGGAATA GACAAAAAGT AAGAAAAACA 15420  
 TGTAATATAT ATATACCAA CAGAGTTCTT CTCTTGTTTG GT 15462

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TTGTCTGGGA AT 12

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

TTGCCTGGGA AT 12

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TTGTTTGGGA AT 12

(2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid

(ii) MOLECULE TYPE: cDNA

TTGTCTGGTA AT

12

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 12 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

AACTTTAAAT TA

12

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 12 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

AACTTAAAT TA

12

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 12 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

TTAAAGACAT TG

12

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 12 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

TTTAAGACAT TG

12

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GCAGATGTCA AG

12

(2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

GCAGATGCCA AG

12

(2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

CGAATCTAAA GA

12

(2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

CGAAGCTAAA GA

12



## (2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

GAAATATTGA TC

12

## (2) INFORMATION FOR SEQ ID NO:59:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GAAACATTGA TC

12

## (2) INFORMATION FOR SEQ ID NO:60:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TCTCTACCCA AC

12

## (2) INFORMATION FOR SEQ ID NO:61:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

TCGTTAACCA AC

12

## (2) INFORMATION FOR SEQ ID NO:62:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid

(ii) MOLECULE TYPE: cDNA

AGTACAATAG GT

12

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 12 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

AGTACTGTGG GT

12

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 12 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

GCACTTGATC CA

12

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 12 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

ACACTGGATC CA

12

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 19 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

CCATCATTGT TGTTGACAA

19

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 19 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

CCATCATTGT GGCTGACAA

19

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 11 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

TTACATGGCC A

11

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 11 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

TCACATGGCG A

11

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 12 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

TTTGGACTGG GC

12

## (2) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

TTTGTATTGG GC

12

## (2) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

GGTCCTAATA CT

12

## (2) INFORMATION FOR SEQ ID NO:73:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

GGGCCTAATA TC

12

## (2) INFORMATION FOR SEQ ID NO:74:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 38 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74

CCATAGAGAG TCCATGGAAA GCGACGCTAA AAACATATC

38

662230 662230 662230